Measuring quality of life: The development and initial validation of the Patient-Reported Erectile Function Assessment instrument

George W. Torrance
McMaster University and *Innovus Research, Inc.*

Margaret-Anne Keresteci
Nexus Research Solutions

Richard Casey
Male Health Centre

Nancy C. Ryan
Innovus Research, Inc.

Jean-Eric Tarride
McMaster University

Objectives: Erectile dysfunction (ED) is a complex condition, which is variously influenced by physical, emotional, societal, and relationship factors. ED has serious implications for the quality of life (QoL) enjoyed by an affected male and his partner. It is very important, therefore, to understand the impact of ED on the QoL of those affected by it. Our objective was to determine if the eight-question Patient Reported Erectile Function Assessment (PREFA) could act as an independent, comprehensive disease-specific instrument in the assessment of QoL as it is impacted by ED.
Methods: During the development and validation of the Erectile Function–Visual Analog Scale (EF-VAS) (14), a new ED-specific preference-based instrument, a series of questions were included at the beginning of the assessment that would act as a way to encourage respondents to focus on their own experience with ED. Upon analysis of the EF-VAS data, it became apparent that the eight-question "warm up" section might act as a stand-alone assessment. Accordingly, the eight questions were named PREFA, and a validation analysis was undertaken to determine their consistency, feasibility, reliability, validity, and responsiveness.
Results: The PREFA questionnaire was found to be feasible and simple to complete, reliable, and valid, with excellent responsiveness. Overall, the PREFA has demonstrated that it can perform as a stand alone, validated assessment of the impact of ED on QoL, assessing areas of QoL not previously captured in existing instruments.
Conclusions: The PREFA is suitable for use in clinical and research settings as a disease-specific QoL assessment tool.

Erectile dysfunction (ED) is defined as the inability to achieve and/or maintain an erection adequate to undertake satisfactory intercourse (9). The most extensive evaluation of ED was performed in the United States in the Massachusetts Male Ageing Study. In this study, 52 percent of men between the ages of 40 and 70 reported some degree of ED and 35 percent of men exhibited moderate to complete ED (6).

Reduced frequency of sexual intercourse is accompanied by feelings of low self-esteem, poor self-image, mental stress, and depression, all of which have a negative impact on the quality of life (QoL) of ED sufferers (1;7). It is important, therefore, that management of patients with ED addresses the psychological aspects of the condition. Within the context of understanding outcomes that are important to patients, ED has presented a challenge to clinicians due to the complex relationships between physical, sexual, and emotional factors. In consideration of these complexities and the combination of psychogenic and organic factors, ED is an area where QoL assessment can be of considerable value in understanding the condition of ED and its treatment.

The EF-VAS (14) is a new ED-specific, self-administered, validated QoL instrument that was designed to quantify the impact of ED on QoL in terms of conventional utility values. As a first step in completing the EF-VAS, the patient answers eight questions about his condition. The purpose of these eight questions is to characterize the extent and the impact of his ED, particularly the impact on his QoL. These eight questions serve to define the patient's self-state for use in the remainder of the EF-VAS. The end product of the EF-VAS is a von Neumann-Morgenstern utility score for the patient's self-state on the conventional health utility scale. The EF-VAS was implemented in a randomized, double-blind, placebo-controlled study of sildenafil citrate conducted in Canada with men diagnosed with ED. The EF-VAS demonstrated that the utility associated with ED was .8 (4). A significant increase to .88 was observed in the sildenafil-treated patients after 12 weeks of treatment.

During the validation analysis of the EF-VAS, it became apparent that the initial eight question section of the instrument may act alone as a useful assessment of QoL in ED. Taken together, the eight questions were brief, comprehensive, and appeared to perform well to assess the impact of ED on QoL. Collectively, the eight questions were named the Patient-Reported Erectile Function Assessment (PREFA). The psychometric and measurement characteristics of the PREFA instrument are reported herein.

METHODS

The data collected in the sildenafil versus placebo study were used to perform the validation analyses of the PREFA. The study included 169 respondents who completed the EF-VAS, which included the PREFA, at screening (Week −1), at baseline (Week 0), and at the end of treatment (Week 12). These individuals were diagnosed with ED and were in a stable relationship at the time of study enrollment. The PREFA was available in English and French Canadian. The validation analyses provide the measurement and psychometric properties of the PREFA for both languages combined. The PREFA consists of eight questions with ordered categorical response options.

To determine whether all eight questions should be included in the final PREFA, the questions were assessed in two steps. First, each question was correlated (Pearson correlation) with every other question. If two questions were highly correlated (>.7), it would indicate that one was possibly redundant and a candidate for exclusion. Second, the internal consistency of the instrument (homogeneity) and the contribution that each question made to the overall instrument was investigated using Cronbach's coefficient alpha, which identifies whether all the questions are measuring the same trait or symptom or whether different traits are being assessed.

Cronbach's coefficient alpha can identify questions that should not be excluded because important differentiation information could be lost. In this case, because the questions in the PREFA were deliberately designed to measure different aspects of the patient's ED, one would not expect or desire maximum homogeneity, and, the interpretation of the coefficient alphas was adjusted accordingly. Cronbach's alpha is a measure of the homogeneity of a scale (greater homogeneity = higher alpha). It measures the extent to which the individual items that constitute the PREFA correlate with one another or with the scale total. To assess homogeneity, a standardized overall Cronbach's alpha coefficient using all the questions was calculated and to which each individual item's standardized Cronbach's alpha was compared. The individual Cronbach's alphas were calculated by excluding each item from the total score and recalculating the coefficient alpha. Correlations between each item and the total score were also calculated. If the item alpha is greater than the standardized overall alpha, it indicates that the removal of that item increases the homogeneity of the scale. That is, the item tends to measure something different from the remaining items. If the item alpha is less than the overall alpha, it indicates that the removal of that item is correlated with the other items in the scale and may be redundant.

Nunnally (10) suggests a value of .7 as an acceptable reliability coefficient, which could be used as a cut point for assessing the item alphas. However, in Proc Corr of SAS version 8.2 a standardized coefficient alpha is calculated and a value of .85 (14) is suggested to identify items for exclusion. Either the cut points of .7 or .85 could be used for

determination of retention or exclusion purposes. The use of the standard prescribed alpha to identify items for exclusion may not be appropriate for this instrument, as the PREFA is not expected to have high internal consistency, because it deliberately measures attributes that are diverse. Instead of the standard alpha, a significant increase or decrease in alpha for this evaluation was defined by the authors as a +/−5 percent change compared with the overall standardized co-efficient alpha. Items that caused the alpha to decrease or increase by more than 5 percent would be considered for exclusion.

Feasibility is the measure of how practical an instrument is, whether the instrument is easy to understand and complete, and if it is acceptable to patients. The feasibility of the PREFA was partially addressed previously because the PREFA questionnaire was a component of the EF-VAS (15). In the analysis of the specific PREFA items, however, there was no mechanism for timing how long administration took, or how difficult respondents perceived completion of the PREFA alone to be. However, the rate of completion and the number of missed or incorrectly answered responses for the PREFA were available for analysis.

Reliability indicates the degree to which an instrument is able to provide the same response under the same conditions for the same patient completing the questionnaire. Reliability of the PREFA was tested for all respondents combined and for stable patients alone. Stable patients were defined as those receiving no treatment for their ED during the assessment period of 1 week and who indicated that there had been no change in their ED between the two administrations of the questionnaire. The determination of stability was based on the question, "Overall has there been any change in your erectile function since your last visit?" If a patient indicated that a change had occurred, they were then asked to indicate the level of change, through a checklist of choices. Reliability was also tested for patients whose disease state changed between the screening to baseline time periods. It was anticipated that the PREFA would be highly reliable, with an intraclass correlation coefficient (ICC) of at least .7. The ICCs were determined using a repeated measures analysis of variance. The variance between patients and within patients for the two visits was calculated. Variance within patients was estimated by the variability between the patient's response at screening and baseline. The variance between patients was estimated by the variability between the patients' responses regardless of visit. The error is the remaining variability not accounted for by the repeated measures or patient variability. The ICC represents the between patient variance divided by the total variance.

Validity is the assessment of whether an instrument is actually measuring what it was designed to measure. Construct validity tests relationships specified in advance, that would be expected if the instrument is measuring what it was designed to measure. For example, if PREFA is measuring the QoL associated with ED, it would be expected to vary systematically with disease severity as measured by other relevant clinical instruments. Specifically, each of the eight questions of the PREFA were correlated with selected questions or domains in other instruments (the International Index of Erectile Function, IIEF) (12), as well specific domains and questions from various QoL questionnaires (SF-12, psychological well-being, MOS family survey, Rosenburg self-esteem scale, erectile distress scale), which were included in the clinical study of the EF-VAS. The total PREFA score was compared with the total score of the Sexual Health Inventory for Men (SHIM) (3;11). It was hypothesized that the PREFA would measure disease severity in a similar manner as the IIEF and the SHIM and have a Pearson's correlation coefficient of >.5, when regressed against the IIEF erectile function domain and the overall SHIM score.

Responsiveness is perhaps the most crucial property for a QoL instrument, from a clinical perspective. As the disease state improves, the PREFA score should increase, and as the disease state worsens, the PREFA score would be expected to decrease. The change from screening to the end of the study for the PREFA was correlated against the change over the same period in the IIEF Questions 3 and 4 and the overall score for the SHIM. It was anticipated a priori that the correlation of the PREFA and the SHIM would be greater than .5. It was expected that the direction of change of the different instruments would be the same and that the magnitude of change would be in proportion to each other.

Response to treatment was assessed through change in ED state defined as a change in SHIM score or a change in both Questions 3 and 4 of the IIEF. A change in disease status should be reflected by a change in the PREFA score, with the impact of active treatment on disease state being larger than with placebo. The changes from screening for the sildenafil and placebo groups were calculated for respondents for whom the disease state changed and for those for whom the disease state had not changed. Results were compared for the PREFA, the SHIM, and Questions 3 and 4 of the IIEF.

A discriminative QoL instrument should also be able to evaluate severity of disease. The distribution of PREFA scores were compared to the levels of the SHIM score and the levels of the IIEF Questions 3 and 4 at baseline and end of study. For the PREFA scores, we used the EF-VAS disease description scores of mild, moderate, and severe ED as initial benchmarks.

The PREFA scores were calculated by assigning each response a numeric value between 1 and 4, where 1 is the worst possible response and 4 is the most positive response. Therefore, a total score using all eight items would range from 8 (the most severely impacted QoL) to 32, which represents no impact on QoL.

When measured by treatment groups, the responsiveness was assessed through the effect size (ES), the standardized response mean (SRM), and the responsiveness statistic (RS). The ES, SRM, and RS were calculated for the PREFA, the IIEF, the SHIM, and the SHIM scores (i.e. severity level). The

ES was calculated by dividing the raw change score over time by the baseline standard deviation. An ES of between .2 and <.5 indicates a small to moderate effect, .5 to .79 indicates a moderate to large effect, and ≥.8 indicates a large effect (5). To measure the sensitivity of the instrument, the SRM was calculated by dividing the raw change score by the standard deviation of the change score, and the RS was calculated by dividing the change score by the standard deviation of those who did not change (stable respondents).

It is also important to determine the minimally clinically important difference (MCID). The MCID can be described as the change in score that is considered to be important by the patient; or the smallest effect that would lead a clinician to recommend a change in therapy for their patients, or the smallest change in a scale score that would be considered a clinical improvement or worsening by clinicians or patients. The MCID of the PREFA was estimated using the methodology described by Samsa and colleagues (13). The ES benchmarks defined by Cohen (5) are small = .2, moderate = .5, and large = .8, and these benchmarks were used to characterize the magnitude of the ES. The MCID was then equated with the minimum ES that classified as a small effect (.20). Therefore, the MCID for the PREFA was hypothesized to be .20 times the standard deviation at baseline of the PREFA score.

RESULTS

When the eight items in the PREFA were examined pair-wise for redundancy, the Pearson correlation coefficients were all less than .7 except for one set. A Pearson correlation coefficient of .72 was found for the pair of items: Q3, *satisfaction with self-image as far as sexuality is concerned*, and Q7, *satisfaction with sexual aspects of relationship*. Accordingly, one of these items could potentially be considered for exclusion.

Further analysis using Cronbach's coefficient alpha, with a threshold of +/−5 percent of the standardized overall Cronbach's alpha (i.e., >.8506 or <.7696) as the cut point for consideration of exclusion was then undertaken. One item met

the criteria for possible exclusion. Q7 had an item alpha of .765, indicating that it was potentially redundant. Reinforcing the above conclusion, Q7 was the item with the largest item correlation. (Value appears in bold in Table 1.) We elected not to exclude it on the grounds that it was close to the line on both tests, and more importantly its item alpha exceeded .7, an absolute threshold suggested by Nunnally (10). These two questions were seen to provide important information on the impact of ED, and it was concluded that their removal would result in the loss of information. Therefore, the psychometric and measurement properties of the PREFA were analyzed using data from all eight questions.

In terms of feasibility, 98 percent of the PREFAs administered were completed appropriately; that is, all questions were answered and one answer was provided for each question. Overall, in all administrations of the instrument (493 administrations × 8 questions each = 3,944 questions) only 11 questions were either missed or were obviously answered incorrectly, giving a very high absolute "no error rate" of 99.7 percent.

With regard to reliability, the ICC between screening and baseline (Week −1 and Week 0) for stable patients was .78, therefore, exceeding the a priori hypothesis that the instrument would have an ICC of at least .70. Table 2 shows the breakdown of this agreement by question.

To assess construct validity, all questionnaires from the screening visit and the end of study visit were used. For the most part, the results for validity were consistent with the author's hypotheses regarding the strength of relationship expected between each question and its comparison data (Table 3). One question, Q6, regarding *impact on everyday activities* was more highly correlated than anticipated with a moderate relationship with the SHIM item regarding *impact of erection problems*. The construct validity results indicate that PREFA does measure what it is intended to measure.

To assess responsiveness, the PREFA scores and particularly the change in PREFA scores between baseline and termination were tested for their correlation with the SHIM scores and the change in SHIM scores. All correlations are positive and significant. Of particular importance, the PREFA

Table 1. Measure of Internal Consistency[a]

Question	PREFA-8	
	Item correlation with total score	Cronbach's alpha
Standardized overall coefficient	—	.8101
Q1 Erectile function	.5057	.7914
Q2 Sexual drive/interest	.5020	.7919
Q3 Self-image	.6332	.7724
Q4 Reaction to sexual situations	.5236	.7888
Q5 Mood/emotions	.4926	.7933
Q6 Impact on everyday activities	.4926	.7933
Q7 Satisfaction with sexual relationship	.6832	**.7648**
Q8 Satisfaction with nonsexual relationship	.3704	.8105

[a] For Q1 to Q8, Chronbach's alpha is calculated with this question removed.
PREFA, Patient Reported Erectile Function Assessment.

Table 2. Reliability: Percentage of Stable Patients with Same Response at Screening and Baseline (Week −1 and Week 0) by Question

	All patients				Stable patients only			
	Same response		Different response		Same response		Different response	
	N	%	N	%	N	%	N	%
Q1 Attain firm long-lasting erections	131	81.4	30	18.6	129	83.8	25	16.2
Q2 Sexual drive/interest	99	60.4	65	39.6	93	59.2	64	40.8
Q3 Self-image	112	68.3	52	31.7	109	69.4	48	30.6
Q4 React to situations	102	63.0	60	37.0	99	63.9	56	36.1
Q5 Mood	114	69.5	50	30.5	108	68.8	49	31.2
Q6 Impact on everyday activities	109	66.9	54	33.1	104	66.7	52	33.3
Q7 Satisfaction with sexual relationship	106	64.6	58	35.4	102	65.0	55	35.0
Q8 Satisfaction with nonsexual relationship	100	61.3	63	38.7	97	62.2	59	37.8

Table 3. Construct Validity: Spearman Correlation of Each Question with Relevant Other Data

	N	Spearman correlation	ASE of Spearman correlation	Hypothesized relationship[a]
Q1 with IIEF erectile function domain	324	.8671	.0170	Strong
Q2 with IIEF sexual desire domain	329	.5562	.0429	Moderate
Q3 with QoLQ self-esteem domain	324	.2272	.0545	Weak
Q4 with IIEF Q6	329	.3846	.0501	Weak
Q5 with QoLQ positive well-being	327	.5565	.0422	Weak
Q6 with QoLQ impact of erection Problems	328	.5329	.0421	Moderate
Q7 with IIEF Q14	328	.6673	.0407	Weak
Q8 with QoLQ, Part III Q7	328	.4345	.0488	Moderate

[a] Strong correlation expected when the constructs are highly similar. Weak correlation expected when there is only a partial overlap of the constructs.
IIEF, International Index of Erectile Function; QoL, quality of life; ASE, asymptotic standard error.

change score for the treatment group is highly correlated with the SHIM change score for the same patients. Thus, if SHIM is a responsive instrument, the PREFA should also demonstrate responsiveness. That the correlation is weaker for the change scores in the placebo group reflects that there was less change in that group; hence, it is more difficult to demonstrate a strong linear correlation between the two measures.

In addition, the following responsiveness statistics were calculated for those patients who experienced a change: ES,

SRM, and RS (Table 4). Two definitions of change were investigated: a patient was defined as experiencing change if his SHIM grade changed, and a patient was defined as experiencing change if his answers to IIEF Q3 or Q4 changed. The PREFA was established to have a large ES, SRM, and RS, regardless of which definition of stable disease status was used and, therefore, demonstrated excellent responsiveness.

The estimated minimum clinically important difference (MCID for the PREFA score was about 1 unit; Table 4). The

Table 4. Responsiveness Indices, Including MCID

	Indices with SHIM grade defining stable patients				Indices with Q3 and Q4 defining stable patients			
	Effect size	Standardized response (mean)	Responsiveness statistic	Minimally important difference	Effect size	Standardized response (mean)	Responsiveness statistic	Minimally important difference
PREFA	1.33	1.02	1.82	.82	1.07	.87	1.44	.82
SHIM	1.54	1.12	4.76	1.09	1.09	.84	2.80	1.12
SHIM grade	1.63	1.10	—		1.12	.78	2.26	
IIEF Q3	.96	.80	1.96		.72	.63	—	
IIEF Q4	1.25	.92	2.38		1.00	.74	—	

MCID, minimally clinically important difference; SHIM, Sexual Health Inventory for Men; PREFA, Patient Reported Erectile Function Assessment; IIEF, International Index of Erectile Function.
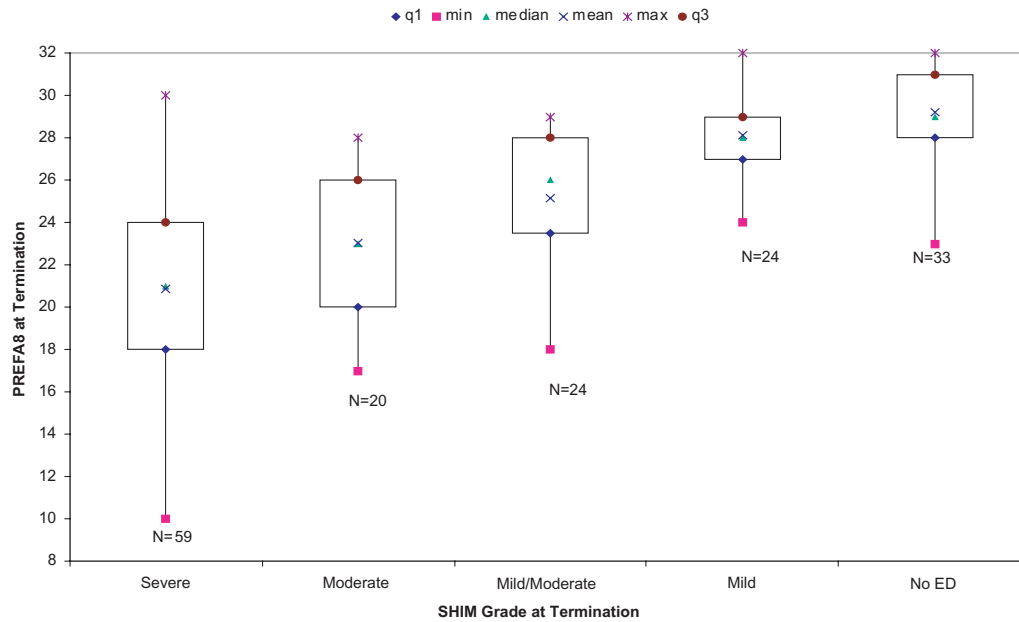
Figure 1. Box and whisker plots of Patient Reported Erectile Function Assessment (PREFA) by Sexual Health Inventory for Men (SHIM) grade at Week 12.

MCID was based on the standard deviation at screening of those respondents who reported a change, combined with the ES benchmark defined a priori as small (.2).

With the treatment groups combined, the PREFA and the SHIM grades were compared at baseline and termination (Figure 1). For men who had an opportunity for sexual activity and intercourse, the SHIM classification for ED is partitioned into five severity grades: no ED (SHIM total score, 22–25), mild ED (17–21), mild to moderate (12–16), moderate (8–11), and severe ED (1–7) (3). Figure 1 presents the distribution of PREFA scores by SHIM score at end of treatment for all patients. As expected, PREFA scores increase systematically as disease severity decreases.

DISCUSSION

The original intent of implementing the eight questions that form the PREFA into the EF-VAS was to prompt the respondents to focus on the diverse aspects of his experience with ED. The observation that these questions seemed capable of independently evaluating the individual's ED led to the hypothesis that the PREFA possessed good psychometric properties. The results of the validation analysis performed on this initial data set were positive and encouraging.

As part of the development of this instrument, a shorter version of the PREFA was explored, with six questions instead of eight. Based on the analysis reported here, however, all eight items are recommended for inclusion in the total score. These analyses pointed to the conclusion that the items surrounding the *impact of sexual difficulties on everyday activities* and *the satisfaction with the nonsexual*

aspects of a relationship actually added valuable information in the total score. These two items are unique to the PREFA and provide information that is not captured in other disease-specific instruments currently used in ED research and clinical management.

Some limitations were identified. The protocol to validate the EF-VAS was not designed to test the PREFA; therefore, the feasibility assessment of the PREFA is incomplete. Time to complete the eight questions was not captured independently; however, the median time to complete the entire EF-VAS was 17 minutes at study end. It is likely that completion of the PREFA would take considerably less than half this time, more likely in the range of 3 to 8 minutes. This short amount of time to complete the PREFA negates any potential concerns regarding fatigue influence on responses and the logistics of administration in clinical practice. In addition, results indicate that the PREFA is very easy to complete, with a low error rate and a high completion rate. This finding suggests that the PREFA is entirely appropriate for use in a clinical setting, with no assistance required for completion.

The defining of a "true" MCID remains a challenge. To assess the MCID appropriately, the question, "*Overall, has there been any change in your erectile function since your last visit*?" with specified levels of response should be included with each administration of the PREFA. This approach will be useful in defining stable disease status in respondents, those who did have a meaningful change, those who had what would be defined as a minimally meaningful change, and those who had what would be defined as no meaningful change.

Recently, several new ED-specific QoL instruments, including the SEAR (self-esteem and relationship) (1;2) and the ED-EQoL (a quality of life measure for patients with erectile dysfunction) (8), have been developed. Further data to test their psychometric properties will be collected in future clinical trials and will add to the evidence to support the validation of these instruments. Future research should be directed at using the PREFA scoring to establish accurate cutoff points for classification of disease severity. If the PREFA proves capable of this, it will add yet another important tool that currently is not available to the clinical assessment mechanisms in the treatment of ED.

This initial validation analysis indicates that the PREFA is a feasible, reliable, valid, and highly responsive instrument that can be incorporated appropriately into either a research setting or a clinical practice setting, without added burden to the respondent or the clinician. The PREFA is easy to complete without assistance in a short amount of time. The reliability analyses indicate that the PREFA is a reliable instrument. The validity analysis indicates that it is measuring what it is intended to measure. The results of the responsiveness analysis were best when the SHIM score was used as the indicator of disease severity. This finding is probably a reflection of the fact that the items in the SHIM and the PREFA are more closely aligned with each other. In fact, responsiveness of the PREFA was excellent and exceeded our expectations. All together, the responsiveness analyses indicate that the PREFA is capable of a sound assessment of change in disease status and, as such, can dependably assess the impact of treatment.

The excellent responsiveness of the PREFA combined with the short, easy to complete format suggest that the PREFA can act as a sound clinical tool, establishing severity of disease, impact on QoL, and the impact of treatment on an individual's ED. Based on this analysis, the benefit seen for the PREFA compared to other instruments is that it addresses various, diverse components of the impact of ED that would otherwise require the administration of several different instruments to capture the same level of information that is captured in this short, easy to administer eight-question format.

CONTACT INFORMATION

George W. Torrance, PhD (Torrance@mcmaster.ca), Professor Emeritus, Department of Clinical Epidemiology and Biostatistics, McMaster University, 25 Main Street W., Hamilton, Ontario L8P 1H1, Canada; Vice President, Department of Scientific Affairs, Innovus Research, Inc., c/o 321 Markland Drive, Toronto, Ontario M9C 1R4, Canada
Margaret-Anne Keresteci (mkeresteci@nexusresearch.ca), Principal, Nexus Research Solutions, 24 Almond Avenue, Thornhill, Ontario L3T 1L1, Canada

Richard Casey, MD (drcasey@malehealth.com), President, Male Health Centre, 407-1235 Trafalgar Road North, Oakville, Ontario L6H 3PI, Canada
Nancy C. Ryan (nryan@innovus.com), Director, Biostatistics and Data Management, Innovus Research, Inc., 1016-A Sutton Drive, Suite 200, Burlington, Ontario L7L 6B8, Canada
Jean-Eric Tarride, PhD (tarride@mcmaster.ca), Assistant Professor, Department of Clinical Epidemiology and Biostatistics, McMaster University, 25 Main Street, W, #2000, Hamilton, Ontario L8P 1H1, Canada

REFERENCES

1. Althof SE. Quality of life and erectile dysfunction. *Urology.* 2002;59:803-810.
2. Cappelleri JC, Althof SE, Siegel RL, et al. Development and validation of the self-esteem and relationship (SEAR) questionnaire in erectile dysfunction. *Int J Impot Res.* 2004;16:30-38.
3. Cappelleri JC, Siegel RL, Glasser DB, et al. Relationship between patient self assessment of erectile dysfunction and the sexual health inventory for men. *Clin Ther.* 2001;23:1707-1719.
4. Casey R, Tarride JE, Keresteci MA, Torrance GW. The Erectile Function Visual Analog Scale (EF-VAS): A disease-specific utility instrument in the assessment of erectile function. *Can J Urol.* 2006;13:3016-25.
5. Cohen J. *Statistical power analysis for the behavioural sciences.* 2nd ed. Hillsdale, NJ: Lawrence Associates; 1988.
6. Feldman HA, Goldstein I, Hatzichristou DG, et al. Impotence and its medical and psychosocial correlates: Results of the Massachusetts Male Aging Study. *J Urol.* 1994;151:54-61.
7. Giuliano F, Pena BM, Mishra A, et al. Efficacy results and quality-of-life measures in men receiving sildenafil citrate for the treatment of erectile dysfunction. *Qual Life Res.* 2001;10:359-369.
8. MacDonagh RP, Porter T, Pontin D, Ewings P. The ED-EQoL: The development of a new quality of life measure for patients with erectile dysfunction. *Qual Life Res.* 2004;13:361-368.
9. National Institute of Health (NIH). Consensus development panel on impotence. *JAMA.* 1993;270:83-90.
10. Nunnally J. *Psychometric theory.* New York: McGraw-Hill; 1978.
11. Rosen RC, Cappelleri JC, Smith MD, et al. Development of an abridged, 5-item version of the International Index of Erectile Function (IIEF-5) as a diagnostic tool for erectile dysfunction. *Int J Impot Res.* 1999;11:319-326.
12. Rosen RC, Riley A, Wagner G, et al. The International Index of Erectile Function (IIEF): A multidimensional scale for assessment of erectile dysfunction. *Urology.* 1997;49:822-830.
13. Samsa G, Edelman D, Rothman ML, et al Determining clinically important differences in health status measures: A general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics.* 1999;15:141-155.
14. SAS OnlineDocV8,1999.
15. Torrance GW, Keresteci MA, Casey R, et al. Development and initial validation of a new preference based disease specific health related quality of life instrument for erectile function. *Qual Life Res.* 2004;13:349-359.