Responsiveness and Minimum Important Differences for the Erection Quality Scale

Raymond C. Rosen,*,† John Wincze, Martin D. Mollen, Kathleen Gondek,‡ Lori D. McLeod§ and William A. Fisher‡

From the Department of Psychiatry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey (RCR), New England Research Institutes, Watertown, Massachusetts (RCR), Departments of Psychology and of Psychiatry and Human Behavior, Brown University, Providence, Rhode Island (JW), Internal Medicine, Arizona Research Center, Inc., Phoenix, Arizona (MDM), Bayer Pharmaceuticals Corporation, West Haven, Connecticut (KG), Research Triangle Institute Health Solutions, Research Triangle Park, North Carolina (LDM), and Departments of Psychology and of Obstetrics and Gynaecology, University of Western Ontario, London, Ontario, Canada (WAF)

Purpose: We evaluated the responsiveness and treatment sensitivity of the Erection Quality Scale, and provided further psychometric validation of this scale.

Materials and Methods: An 8-week, placebo controlled, randomized clinical trial investigating the efficacy and safety of vardenafil in patients with erectile dysfunction was performed. The Erection Quality Scale, together with a number of other patient and partner questionnaires, was administered at a screening visit, at baseline, and weeks 4 and 8 of treatment. Erection Quality Scale responsiveness was investigated by evaluating treatment induced changes and modeling using ANCOVA. Internal consistency, convergent and discriminant validity, and minimum important difference of the Erection Quality Scale were also assessed.

Results: Efficacy evaluations demonstrated that the Erection Quality Scale was sufficiently responsive to differentiate the treatment benefits of vardenafil compared with placebo. Internal consistency for the Erection Quality Scale total score was similar across visits, with values high enough to suggest reliability of items included in the scale. Discriminant validity of the Erection Quality Scale total score was demonstrated, with a high correlation with the erectile function domain of the International Index of Erectile Function (0.88, $p < 0.0001$) and negligible correlations with clinical measures assumed to be unrelated to erection quality. All Erection Quality Scale total score comparisons substantially exceeded the 5-point minimum important difference estimate.

Conclusions: The Erection Quality Scale was responsive and internally consistent, and demonstrated convergent and discriminant validity. Furthermore, this instrument provided a unique contribution to the measurement of erection quality compared to the International Index of Erectile Function. This study provides strong evidence supporting the use of the Erection Quality Scale in clinical trials.

Key Words: impotence, vardenafil, questionnaires

Several patient based assessments have been developed to investigate the treatment of ED from the patient perspective. These have primarily focused on functional aspects such as time to onset of treatment effect, duration of action, and ability to achieve and maintain an erection sufficient for sexual intercourse (eg International Index of Erectile Function, sexual encounter profile diary).[1] Other instruments such as the Self-Esteem And Relationship Questionnaire measure the impact of ED on aspects of quality of life.[2] Finally, treatment satisfaction instruments make more direct treatment evaluations from patient and partner perspectives (eg Treatment Satisfaction Scale).[3]

Aspects such as amount of stimulation required to achieve erection, degree of firmness or rigidity achieved, duration of erection and sensitivity of the penis are not adequately addressed with existing instruments. To better assess these dimensions the 15-item patient completed Erection Quality Scale was developed.[4] The sum of the responses to the 15 items yields an overall score and higher scores indicate a more favorable erection quality. The EQS has been validated previously in a 4 site, 200 patient, prospective, test-retest study.[4] This report describes the first evaluation of EQS responsiveness and treatment sensitivity in a placebo controlled, randomized clinical trial, and provides further psychometric validation including initial estimates of the minimum important difference.

MATERIALS AND METHODS

Study Design

The EQS was evaluated in a randomized, double-blind, placebo controlled, parallel group, multicenter United States

Submitted for publication February 26, 2007.
Supported by Bayer Pharmaceuticals Corporation.
* Correspondence: New England Research Institutes, 9 Galen St., Watertown, Massachusetts 02472 (telephone: 617-923-7747 x383; FAX: 617-926-8246; e-mail: rrosen@neriscience.com).
† Financial interest and/or other relationship with Pfizer, Bayer and Sanofi-Aventis.
‡ Financial interest and/or other relationship with Bayer Pharmaceuticals.
§ Financial interest and/or other relationship with Research Triangle Institute Health Solutions and Bayer.

DOI:10.1016/j.juro.2007.07.019

study of vardenafil (vardenafil HCl, Levitra®) vs placebo in 105 and 111 men, respectively, with ED. Patients were randomized to therapy using a 1:1 ratio via a computer generated random code. The methods of this trial have been described in detail elsewhere.[5] Following a 4-week untreated baseline period, patients received 10 mg per dose of vardenafil or placebo for a 4-week double-blind treatment period. During the next 4 weeks patients could titrate up to 20 mg or down to 5 mg per dose of vardenafil, or matching double-blind placebo for a second double-blind treatment period. A final evaluation (visit 4) was made within 24 hours of the last dose of the study drug.

The EQS, IIEF, SEP, GAQ and a single-item global rating of change were administered at screening visit 1 (week −4), baseline visit 2 (week 0), visit 3 (week 4) and visit 4 (week 8). The EF domain of the IIEF provides scores ranging from 1 to 30 with higher scores indicative of better function. The SEP2 and SEP3 scores are the proportion of subjects able to insert the penis into the partner's vagina or having an erection lasting long enough for successful intercourse, respectively. The GAQ score is the percentage of subjects indicating an improvement in erections in the last 4 weeks compared with before the study. The global rating of change assessed how patient erections changed in the last 4 weeks using a 5-point Likert scale ranging from much better to much worse. This single-item scale provided an anchor for the MID evaluation.

Study Population

The study population included men 18 years old or older who had ED for 6 or more months and who were currently in a heterosexual relationship. Full details of the inclusion and exclusion criteria have been published elsewhere.[5]

Efficacy and Safety Assessments

The primary efficacy measure was EQS responsiveness, assessed as the change in total score between baseline (visit 2) and 2 months after treatment (visit 4). To account for dropouts (5% between visits 3 and 4), analyses of the primary efficacy measure used the LOCF method.

Secondary efficacy measures included EQS total score change comparisons between vardenafil and placebo from baseline at visit 2 to visits 3 and 4 of treatment, the IIEF-EF score change comparisons at visit 4 or LOCF, and per patient overall success rate comparisons at visit 4 or LOCF as reported by SEP2 and SEP3.

Each efficacy evaluation used an ANCOVA with baseline scores as the covariate, and treatment group and investigational center included as main effects in the model. Results were reported as least squares means and all tests were 2-sided. Safety was evaluated in terms of premature termination, adverse events, concomitant medication use and changes from baseline in all vital signs. The LOCF method was used unless otherwise noted.

Responsiveness

Responsiveness was assessed by evaluating treatment differences as defined by the primary efficacy outcome (EQS responsiveness) and the first secondary outcome (EQS change week 0 to week 4). Specifically the change in EQS total score from baseline at visit 2 to visits 3 and 4 were modeled using an ANCOVA. Effect size statistics were then calculated using 2 approaches. The first estimated a group-level effect size, defined for each treatment group as the average change from baseline (visit 2 to visit 4 or LOCF) score for that group (either vardenafil or placebo) divided by the standard deviation of the visit 2 score for that same group. The second approach used a modified Guyatt's responsiveness statistic, defined as the difference in average change from baseline (visit 2 to visit 4 or LOCF) scores between the 2 treatment groups divided by the standard deviation of change scores in the placebo group. These statistics were also computed using visit 3 and visit 4 data for comparison. Effect sizes of 0.2, 0.5 and greater than 0.8 were considered small, moderate and large, respectively.[6]

Reliability and Validity

Cronbach's α was used to assess the internal consistency of the EQS. Additional evidence for the convergent and discriminant validity of the instrument was gathered by correlating EQS total change scores with change on a range of clinical measures, some of which are theoretically related to erection quality (eg IIEF-EF, SEP2, SEP3) while others are not (eg heart rate, height). Implicit in the examination of the relationships between the EQS total scores and the 3 measures of erectile function is the assumption that changes in erection quality and erectile function are related. Gathering support for convergent validity then involves demonstrating that the EQS measures erection quality precisely enough to demonstrate this relationship. Because positive values represent an increase in erection quality as measured by the EQS and less severity on the 3 related clinical measures, it was expected that any correlation between these measures would be positive. Moreover, it was expected that these correlations would be moderate to large because the constructs addressed by each measure, while related, are not redundant. Correlations between the IIEF-EF and SEP2 and SEP3 were also computed to evaluate the differences between the IIEF-EF and the EQS. In addition, a series of models were computed to further compare the constructs measured by these 2 instruments. Specifically 3 models tested the predictive value of the IIEF-EF and EQS total scores for predicting SEP2 and SEP3. A regression model was constructed using SEP2 change score from baseline (percent) as the dependent variable, and age, EQS total change score from baseline, IIEF-EF domain score change from baseline and treatment (dichotomous, vardenafil = 1 and placebo = 0) as predictors. Reduced models were then constructed without either EQS change score or IIEF-EF domain change score. To assess the responsiveness of the EQS change score and IIEF-EF domain change score in relation to the SEP2 and SEP3 percent change, the following variables were log(10) transformed: age, EQS change score and IIEF-EF domain change score. A full model with transformed variables was then reconstructed.

Finally, a model was calculated to predict GAQ scores based on age, treatment, IIEF-EF change score and EQS change score. Together these models provided an evaluation of the relationship between erection quality as measured by the EQS and IIEF-EF as predictors of key clinical variables (SEP2 and SEP3) as well as the global rating of change (GAQ).

A known-groups analysis comparing responder subgroups was also conducted to provide further support for the convergent and discriminant validity of the EQS. For this