Psychometric Properties of the Male Urogenital Distress Inventory (MUDI) and Male Urinary Symptom Impact Questionnaire (MUSIQ) in Patients Following Radical Prostatectomy

Joanne P. Robinson
Tamara Avi-Itzhak
Ruth McCorkle

The Male Urogenital Distress Inventory (MUDI) and Male Urinary Symptom Impact Questionnaire (MUSIQ) were developed to provide a measure of the specific impact of lower urinary tract symptoms (LUTS) on health-related quality of life in community-dwelling men. The MUDI and MUSIQ are based on the widely used Urogenital Distress Inventory and Incontinence Impact Questionnaire (Shumaker, Wyman, Uebersax, McClish, & Fantl, 1994), which were designed for community-dwelling women. Use of the MUDI and MUSIQ may provide more precise measurement of health-related quality of life in men with LUTS. Initial testing of the MUDI and MUSIQ in a racially and educationally diverse sample of men with LUTS yielded evidence of internal consistency reliability, content validity, construct validity, and sensitivity of both instruments were evaluated in a cohort of men who participated in a larger study of nursing’s impact on quality of life post-prostatectomy. Findings support the reliability, validity, and sensitivity of both instruments in this population.

Introduction
Research instruments can be used to assess the impact of lower urinary tract symptoms on health-related quality of life. Evaluating the reliability, validity, and sensitivity of two of these instruments provides evidence concerning their readiness for use in research and clinical settings.

Objective
The researchers evaluated psychometric properties of the Male Urogenital Distress Inventory and Male Urinary Symptom Impact Questionnaire in men suffering from lower urinary tract symptoms following radical prostatectomy.

Method
The internal consistency reliability, concurrent validity, construct validity, and sensitivity of both instruments were evaluated in a cohort of men who participated in a larger study of nursing’s impact on quality of life post-prostatectomy.

Results
The reliability, validity, and sensitivity of both instruments were supported in this population.

Conclusions
Initial support is provided for use of the instruments to measure the effect of interventions for lower urinary tract symptoms and urine leakage on health-related quality of life during the first 6 months of recovery from radical prostatectomy; however, further research with a larger and more diverse sample is needed prior to using the instruments in clinical settings.

Authors’ Note: Funding Sources include the American Cancer Society (TBP–98–118–01–PBP) and Rutgers College of Nursing Faculty Research Development Award.
current validity, and construct validity (Robinson & Shea, 2002). The purpose of this study was to examine reliability, validity, and sensitivity of the MUDI and MUSIQ in a sample of men with LUTS following radical prostatectomy for prostate cancer.

BACKGROUND

Health-Related Quality of Life

Health-related quality of life (HRQL) is a multidimensional construct conceived by clinicians, health service researchers, and policymakers to explore aspects of quality of life associated with health and health care (Guyatt, Feeny, & Patrick, 1993; Haas 1999). HRQL pertains to an individual’s appraisal of their sense of well-being; the extent to which reasonable physical, emotional, and intellectual function are maintained; and the degree to which the ability to participate in valued activities within the family, workplace, and community is retained (Donovan et al., 2005; Naughton & Shumaker, 1998).

Primary dimensions of HRQL include physical functioning, psychological functioning, social and role functioning, overall life satisfaction, and perceived health status. Secondary dimensions, such as cognitive or neuropsychological functioning, personal productivity, intimacy, sexual functioning, sleep disturbance, pain, illness-specific symptoms, and adverse effects of treatment are also often measured depending on the context in which HRQL is examined (Berzon, Hays, & Shumaker, 1993; Naughton & Shumaker, 1998). Clinicians, researchers, and policymakers rely on information concerning HRQL for determining best practices, evaluating treatment alternatives, addressing treatment adherence, measuring care quality and clinical effectiveness, and making decisions about allocation of resources and reimbursement.

Instruments that measure HRQL are either generic or condition specific. Generic instruments are designed to assess HRQL across a broad range of health conditions or populations, allowing for comparisons across diverse conditions or groups. Generic instruments may, however, lack sensitivity to the unique impact of specific health problems and efficacy of specific treatments. In contrast, condition-specific instruments are designed to assess HRQL in particular diseases, conditions, age groups, or ethnic groups. Although use of condition-specific instruments restricts comparison of HRQL across diverse health conditions or groups, greater sensitivity to clinically relevant changes in specific conditions over time is usually achieved (Patrick & Deyo, 1989; Ware, 1994). Thus, condition-specific measures are generally superior for capturing the unique impact of conditions like LUTS on HRQL, as well as the efficacy of interventions for such conditions from the patient’s perspective.

The MUDI and MUSIQ

The MUDI and MUSIQ are companion instruments that together provide a condition-specific measurement of HRQL for men with LUTS. Items were derived respectively from the Urogenital Distress Inventory and Incontinence Impact Questionnaire (Shumaker et al., 1994) and adapted through interviews with male urology patients and content experts (Robinson & Shea, 2002). The MUDI is a 27-item questionnaire that prompts respondents to indicate if they experience various LUTS and the degree to which each symptom bothers them. Responses are recorded on a 5-point scale ranging from 1 (symptom not present) to 5 (symptom present and greatly bothersome). The MUSIQ is a 32-item questionnaire that prompts respondents to indicate if they experience various LUTS and the degree to which each symptom bothers them. Responses are recorded on a 4-point scale ranging from 0 (not at all) to 3 (greatly).

In initial testing with male urology clinic patients (Robinson & Shea, 2002), the MUDI and MUSIQ demonstrated internal consistency reliability, indicating that all items were homogeneous or connected to a single construct (Streiner & Norman, 1995). Content validity was also demonstrated, suggesting that both instruments covered the universe of items related to HRQL in men with LUTS adequately. The moderate correlation observed between MUDI and MUSIQ scores provided preliminary evidence of concurrent validity, the extent to which a relationship exists between scores on a target instrument and scores on an accepted measure of the same construct administered simultaneously. Finally, construct validity, the extent to which an instrument actually measures the construct under investigation, was supported in two ways. First, MUDI and MUSIQ scores were related to sociodemographic and self-reported co-morbidity data reflective of HRQL that were gathered simultaneously. Second, fundamental dimensions of HRQL were represented in the factor structures or subscales of the instruments.

METHODS

In this study, internal consistency reliability, concurrent validity, construct validity, and sensitivity of the MUDI and MUSIQ were evaluated in a cohort of men who underwent radical prostatectomy for newly diagnosed prostate cancer. These men were participants in a larger study that examined nursing’s impact on quality of life post-prostatectomy (McCorkle, Pickett, Malkowicz, & Robinson, 1998; McCorkle, Sieffert, Dowd, Robinson, & Pickett, 2007).

Sample

The sample for the parent study was recruited from the urology practices of two academic medical centers and limited to patients who were English-speaking, newly diagnosed with prostate cancer, married or in a committed relationship, elected radical prostatectomy as primary treatment, and resided within a 50-mile radius of the study cen-
ter. Recruitment of participants was initiated after approval to conduct the study was obtained from the institutional review boards of each institution. The sample for the study reported herein consisted of the first 50 participants in the intervention group and the first 50 participants in the control group who met the following criteria: (a) completed a baseline MUDI and MUSIQ at the end of the first postoperative month, and (b) remained enrolled in the study through at least the third postoperative month. Thus, the sample consisted of 100 men (50 participants from each group). We oversampled by 10% to accommodate the possibility of missing data. The study had 80% power to detect correlations of 0.30 and medium-to-large differences between the means of two groups, assuming a Type I error rate of 0.05.

Procedure

Recruitment of participants into the parent study occurred at individual and group pre-operative education sessions conducted by nurses in the urology practices of the participating medical centers. Eligible patients were randomly assigned to intervention or control groups prior to surgery following verbal consent to participate. Written informed consent was obtained during hospitalization on the second postoperative day, after which demographic data were collected and participants were advised of their group assignment. In addition to the routine postoperative care provided during visits to the urologist’s office, members of the intervention group received a series of eight weekly home visits and followup telephone calls from an advanced practice nurse. The intervention commenced upon discharge from the hospital and was designed to facilitate postoperative recovery, assist the patient/spouse dyad with management of symptoms after discharge from the hospital, and promote the couple’s transition from the acute to the chronic phase of the illness (Rolland, 1987).

Data Collection and Instruments

All participants in the parent study completed the MUDI and MUSIQ at 1, 3, and 6 months following surgery, and the SF-36 Health Survey (Ware, Snow, Kosinski, & Gandek, 1993) at 3 and 6 months following surgery. Questionnaires were administered by research assistants either face to face in the participant’s home, or by telephone interview.

In addition, a 24-hour pad test to detect urine leakage was completed by each participant in the intervention group at the conclusion of the first postoperative month. The participant’s advanced practice nurse provided instruction and materials for self-administration of the pad test, and arranged for return of test materials to the study center within 72 hours of test completion.

MUDI. A general description of the MUDI is presented in the background section of this report. In the parent study, values on the original 5-point response scale were changed to range from 0 (symptom not present) to 4 (symptom present and greatly bothersome) to facilitate calculation of the number of symptoms experienced by the subject. These values were retained for the study reported herein. Thus, possible scores on the 27-item MUDI in this study range from 0 (no symptoms and no bother) to 108 (maximum symptoms and maximum bother).

In preliminary testing with a racially and educationally mixed group of male urology clinic patients (Robinson & Shea, 2002), the MUSIQ demonstrated internal consistency reliability with a Cronbach’s alpha coefficient of 0.95. Concurrent validity was supported based on moderate correlation with the MUDI (r=0.59; p<0.001). Construct validity was supported based on significant and appropriate variation of MUSIQ scores with age, desire for socialization, urine leakage, and depression. In addition, principal components analysis with varimax rotation yielded six factors, which accounted for 72.9% of the variance in MUSIQ scores. These factors reflected primary dimensions of HRQL, including physical functioning, psychological functioning, social functioning, and role activities.

SF-36 Health Survey. The SF-36 Health Survey is a generic measure of HRQL that has been used extensively in many populations and is considered comprehensive, psychometrically sound, and efficient (Ware et al., 1993). The instrument consists of 36 items grouped into eight multi-item subscales of two to ten items each, and a single-item measure of change in health. Subscales assess the domains of physical functioning, role limitations due to physical health, bodily pain, general health, vitality, social

max rotation yielded seven factors corresponding to categories of LUTS, a secondary dimension of HRQL. These factors accounted for 65.4% of the variance in MUDI scores.

MUSIQ. A general description of the MUSIQ is presented in the background section of this report. In the parent study, values on the original 4-point response scale, which ranged from 0 (not at all) to 3 (greatly), were changed to range from 1 (not at all) to 4 (greatly) in the interest of consistency with the MUDI. These values were retained for the study reported herein. Thus, possible scores on the 32-item MUSIQ in this study range from 32 (no impact) to 128 (maximum impact).

In preliminary testing with a racially and educationally mixed group of male urology clinic patients (Robinson & Shea, 2002), the MUSIQ demonstrated internal consistency reliability with a Cronbach’s alpha coefficient of 0.95. Concurrent validity was supported based on moderate correlation with the MUDI (r=0.59; p<0.001). Construct validity was supported based on significant and appropriate variation of MUSIQ scores with age, desire for socialization, urine leakage, and depression. In addition, principal components analysis with varimax rotation yielded six factors, which accounted for 72.9% of the variance in MUSIQ scores. These factors reflected primary dimensions of HRQL, including physical functioning, psychological functioning, social functioning, and role activities.

SF-36 Health Survey. The SF-36 Health Survey is a generic measure of HRQL that has been used extensively in many populations and is considered comprehensive, psychometrically sound, and efficient (Ware et al., 1993). The instrument consists of 36 items grouped into eight multi-item subscales of two to ten items each, and a single-item measure of change in health. Subscales assess the domains of physical functioning, role limitations due to physical health, bodily pain, general health, vitality, social
functioning, role limitations due to emotional health, and mental health. Raw scores for each subscale and the health transition item are transformed into standardized scores ranging from 0 (worst) to 100 (best). In this study, however, raw scores were converted to Z scores for comparison with Z scores calculated for the MUDI and MUSIQ since standardized scores for the MUDI and MUSIQ are not available.

24-hour pad test. The 24-hour pad test is a standard objective indicator of the amount or severity of involuntary urine loss. The test produces results that are reproducible, correspond to urodynamic findings of continence or incontinence, and correlate highly with pad testing carried out over a longer period (Groutz et al., 2000; Karantantis et al., 2005; Versi et al., 1996; Wilson, Mason, Herbison, & Sutherst, 1989). To perform the test, patients wear a set of 12 pre-weighed absorbent pads for 2 hours each over a 24-hour period. Used pads are weighed on a gram scale to detect the difference between dry and wet weight, which corresponds to the amount of urine loss in milliliters.

In the parent study, participants in the intervention group were given a set of 12 pre-weighed absorbent pads in individual zip-lock bags labeled from 1 to 12. Pads were worn in sequential order for a maximum of 2 hours while awake. Participants replaced used pads in their original zip-lock bag and labeled the bag with wearing time and main activity(s) during wearing time. Used pads were weighed and returned to the research team at the end of the study period. Each patient was weighed at baseline and provided weight loss data over 72 hours (Versi et al., 1996; Wilson et al., 1989). In this study, however, raw scores were replaced with a “0” if presence of the symptom could not be determined; otherwise, the mean value of the item for all subjects who experienced the symptom in the time period was imputed. Missing values in eligible MUSIQs and SF-36 surveys were replaced with the mean value of the item for all subjects in the time period.

Internal consistency reliability, concurrent validity, construct validity, and sensitivity of MUDI and MUSIQ scores were then examined using the SPSS software package version 9.0. The SF-36 Health Survey served as the standard measure of HRQL for assessment of concurrent validity. The 24-hour pad test served as the objective criterion of involuntary urine loss, which should be related to MUDI and MUSIQ scores if the MUDI and MUSIQ are valid condition-specific co-measures of the construct of HRQL in men with continence issues related to LUTS.

RESULTS

Sample

Data from 14 of the initial 100 participants in the parent study were excluded due to either a missed MUDI and MUSIQ at baseline (n=5) or withdrawal from the parent study prior to the third postoperative month (n=9). These data were replaced with data from the next 14 consecutive enrollees in the parent study who met this study’s inclusion criteria. There were no demographic differences between the participants excluded or the participants who replaced them. Demographic characteristics of the sample are presented in Table 1. Participants ranged in age from 42 to 72 years (M=59.6; SD=6.2) and had from 4 to 21 years of education (M=15.6; SD=3.2). The majority were Caucasian (89%), married (93%), and employed full time (67%). Half reported annual incomes of $90,000 or more. Pre-operative urinary incontinence was denied by the majority of participants (89%).

MUDIs were completed by 100, 87, and 74 participants respectively at 1, 3, and 6 months following surgery (see Table 1). All completed MUDIs contained sufficient data for inclusion in the analyses. MUDI scores demonstrated variability at each time period and declined over time, which is consistent with the expected course of lower urinary tract recovery following radical prostatectomy.

MUSIQs were completed by 100, 87, and 74 participants respectively at 1, 3, and 6 months following surgery (see Table 1). Of these, 1, 5, and 6 MUSIQs respectively at 1, 3, and 6 months following surgery contained insufficient data for inclusion in the analyses. Thus, 99, 82, and 68 MUSIQs respectively from 1, 3, and 6 months following surgery were analyzed in this study. Similar to the MUDI, MUSIQ scores demonstrated variability at each time period and declined over time, which resonates with the pattern of postoperative lower urinary tract recovery depicted by the MUDI.

Reliability

Both instruments demonstrated internal consistency or homogeneity. Cronbach’s alpha coefficients for the MUDI were 0.85, 0.89, and 0.91 respectively at 1, 3, and 6 months following surgery. Cronbach’s alpha coefficients for the MUSIQ were 0.94, 0.97, and 0.97 respectively at 1, 3, and 6 months following surgery.

Concurrent Validity

Concurrent validity was assessed initially by determining the correlation between MUDI and MUSIQ scores at each data collection point. Since the MUDI and MUSIQ measure different dimensions of the same construct, moderate correlations between MUDI and MUSIQ scores were anticipated. As expected, the Pearson’s correlation coefficient between the MUDI and MUSIQ was moderate (r=0.65; p<0.01) at 1 month following surgery; however, the correlation coefficient at both 3 and 6 months was lower (r=0.30; p<0.01) at 3 months following surgery. At 6 months following surgery, the correlation coefficient was moderate (r=0.62; p<0.01).
months following surgery was high \( (r=0.84; p<0.01) \).

Concurrent validity was further assessed by correlating MUDI and MUSIQ scores with scores on the SF-36 Health Survey, an established generic measure of HRQL that was administered simultaneously at the 3 and 6-month postoperative data collection points. Paired MUDI and SF-36 scores were available for 87 and 70 subjects respectively at 3 months and 6 months following surgery. Paired MUSIQ and SF-36 scores were available for 82 and 65 subjects respectively at 3 months and 6 months following surgery. Since the MUDI and MUSIQ are condition-specific HRQL measures, moderate correlations with the SF-36, a generic HRQL measure, were anticipated. As expected, Pearson’s correlation coefficients between the SF-36 and both instruments were in the moderate range at data collection that occurred during postoperative month 3 (MUDI: \( r = 0.29; p < 0.01 \); MUSIQ: \( r = 0.30; p < 0.01 \)) and postoperative month 6 (MUDI: \( r = 0.44; p < 0.01 \); MUSIQ: \( r = 0.65; p < 0.01 \)).

### Construct Validity

Construct validity was evaluated by correlating each participant’s MUDI and MUSIQ scores with the amount of urine leakage experienced by the participant during a 24-hour period. Urine leakage was calculated from results of the 24-hour pad test, which was completed by 42 of the 50 intervention group participants at 1 month following surgery. Leakage volumes ranged from 0 to 1,825 ml \( (M = 404 \text{ ml}; \text{SD} = 422 \text{ ml}) \). Since it is reasonable to assume that the severity of urine leakage influences HRQL, significant correlations between this objective indicator of the severity of urine leakage and scores on the MUDI and MUSIQ were anticipated. As expected, the Pearson’s correlation coefficient between amount of urine leakage and scores on the MUSIQ was 0.44 \( (p < 0.01) \). A moderate correlation between amount of urine leakage and scores on the MUDI was also observed; however, statistical significance was not reached \( (r = 0.30; p = 0.08) \).

### Sensitivity

The ability of the MUDI and MUSIQ to detect clinical change was evaluated by performing paired \( t \)-tests to examine within-subject changes in mean scores from 1 to 3 months, 3 to 6 months, and 1 to 6 months following surgery. Gradual resolution of LUTS typically occurs over the course of 1 year following radical prostatectomy (Robinson, 2006). Thus, declines in both MUDI and MUSIQ scores over time were anticipated, since symptom distress and HRQL impact would be expected to decline in the presence of a decline in LUTS. As
shown in Table 2, both instruments detected statistically significant declines in symptom distress and HRQL impact that occurred from 1 to 3 months and 1 to 6 months following surgery. Although mean MUDI and MUSIQ scores also declined from 3 to 6 months following surgery, the mean change detected was smaller and statistical significance was not reached.

To further evaluate sensitivity to change, effect sizes were calculated to determine the magnitude and meaning of changes in MUDI and MUSIQ scores that occurred from 1 month following surgery (early stage of lower urinary tract recovery) to 6 months following surgery (mid-to-late stage of lower urinary tract recovery). Effect size values are generally classified as small (0.20), medium (0.50), or large (0.80 or higher) and as such provide a basis for determining whether change in a variable is clinically relevant (Cohen, 1987). For this study, effect size was defined as the difference between postoperative month 1 and postoperative month 6 mean values for a variable divided by the standard deviation of the same variable at postoperative month 1 (Kazis, Anderson, & Meenan, 1989). A medium change (0.53) in symptom distress from 1 to 6 months following surgery was detected by the MUDI, and a medium-to-large change (0.72) in HRQL impact from 1 to 6 months following surgery was detected by the MUSIQ.

### DISCUSSION

The results of this study support the reliability, validity, and sensitivity of MUDI and MUSIQ scores as indicators of condition-specific HRQL in men with LUTS following radical prostatectomy for prostate cancer. Both instruments demonstrated a high level of internal consistency. Concurrent validity findings suggest that, as expected, relationships exist between MUDI and MUSIQ scores at all data collection points. Additional evidence of concurrent validity consists of the relationships between the SF-36, a widely accepted generic measure of HRQL, and both instruments at 3 and 6 months following surgery.

Construct validity findings indicate that a positive relationship exists between scores on the MUSIQ and severity of urine leakage, which were both evaluated at 1 month following surgery. The relationship between MUDI and MUSIQ scores and severity of urine leakage was in the expected direction, but weaker and did not reach statistical significance. This may have occurred because the MUDI addresses a broad array of LUTS in addition to urine leakage. Finally, there is evidence that both instruments are sensitive to expected changes in symptom distress and HRQL impact. Significant declines in both MUDI and MUSIQ scores were observed as expected over the first 6 postoperative months, which reflected changes in symptom distress and HRQL impact that occurred over this time period.

Findings from this study suggest two possible shortcomings of the instruments. First, the high correlation between MUDI and MUSIQ scores at 3 and 6 months following radical prostatectomy may indicate unnecessary conceptual overlap between the instruments. Second, there is no evidence that either instrument demonstrates the sensitivity to detect small clinical changes that might be important to track in individual patients in order to make decisions regarding treatments or their evaluation. In addition, the study is limited by the relative homogeneity of the
sample in terms of race, marital status, education, income, and geographic location. Additional testing of the MUDI and MUSIQ with a larger and more diverse sample of radical prostatectomy patients is recommended to further evaluate sensitivity, validate existing subscales, explore item reduction, and assess test-retest reliability. This study extends the uses of the MUDI and MUSIQ to an important population affected by LUTS and urine leakage. Initial support is provided for use of the instruments to measure the effect of interventions for LUTS and urine leakage on HRQL in men during the first 6 months of recovery from radical prostatectomy for prostate cancer; however, further research is needed prior to adopting the MUDI and MUSIQ as measures in the clinical setting.

References