Validation of the lower urinary tract symptom score

Jerry G. Blaivas, MD,[1,2,3] Johnson F. Tsui, MD,[2,4] Gabriel Mekel, MD,[2,5] Matthew S. Benedon, BA,[2] Belinda Li, MD,[6] Fara M. Friedman, MD,[3] James M. Weinberger, BS,[2,7] Jeremy Weedon, PhD,[3] Jeffrey P. Weiss, MD[2,3]

[1]Weill Cornell College of Medicine, New York, New York, USA / [2]Institute for Bladder and Prostate Research, New York, New York, USA
[3]SUNY Downstate Medical Center, Brooklyn, New York, USA / [4]North Shore-LIJ Lenox Hill Hospital, New York, New York, USA
[5]Drexel University College of Medicine, Philadelphia, Pennsylvania, USA
[6]Loyola University Chicago Strictch School of Medicine, Chicago, Illinois, USA
[7]David Geffen School of Medicine at UCLA, Los Angeles, California, USA

BLAIVAS JG, TSUI JF, MEKEL G, BENEDON MS, LI B, FRIEDMAN FM, WEINBERGER JM, WEEDON J, WEISS JP. Validation of the lower urinary tract symptom score. *Can J Urol* 2015;22(5):7952-7958.

Introduction: *To develop and validate a lower urinary tract symptom score (LUTSS) as a measure of lower urinary tract symptom (LUTS) severity and a treatment outcome tool in adults.*

Materials and methods: *An expert panel was convened to develop the LUTSS questionnaire. Content validity was achieved by obtaining subject and expert feedback from two prospective drafts. Subjects were divided into three groups: normal, LUTS and overactive bladder (OAB). Questionnaire was administered on two separate occasions within 1-2 weeks. Test-retest reliability, internal consistency, discriminant validity, criterion validity and responsiveness to change were also assessed.*

Results: *The questionnaire contains 14 questions with answers scored on a 5-point Likert scale (0-4). It includes 9 storage, 4 voiding and 1 bother question. One hundred ninety-one patients completed it; 80 males and 111 females, mean age 65 years (range 22-91). Seventy-two had OAB, 91 LUTS without OAB and 28 were normal. Test-retest intraclass correlation was 0.96 and Cronbach's-α was 0.77, indicating strong test-retest reliability and internal consistency, respectively. ANOVA and post-hoc bootstrap-generated adjustments showed significant differences between the three groups ($p < 0.001$), demonstrating discriminant validity. Responsiveness to change was exhibited by the significant decrease between preop and postop scores and a concurrent patient global impression of improvement (PGI-I) score indicative of symptomatic improvement.*

Conclusion: *The 14-question LUTSS is a validated questionnaire that assesses a full range of LUTS in men and women. The ordinal nature of the data with its highly specific description of symptoms makes it ideally suited as a nuanced and comprehensive symptom score and patient reported outcome (PRO) tool.*

Key Words: patient reported outcomes (PRO), lower urinary tract symptoms, questionnaire, voiding dysfunction, overactive bladder, urinary incontinence

Introduction

Diagnostic and treatment outcomes for lower urinary tract disorders are predicated on a valid assessment of lower urinary tract symptoms (LUTS).[1] According to the International Consultation on Incontinence (ICI) and FDA guidelines, symptom assessment should be accomplished with psychometrically validated patient reported outcome (PRO) tools. Over 70 such questionnaires have been validated and the ICI has recommended that existing questionnaires should be utilized for clinical research rather than developing new PRO's.[2] Despite this plethora of PRO's and despite the ICI's recommendation, in our judgment, all existing questionnaires have sufficient deficiencies to warrant consideration of yet another one.[3-12] Examples of questionnaires include the American Urological Association Symptom Score (AUASS),[3] Overactive Bladder Symptom Score (OABSS),[13] International Consultation on Incontinence Modular Questionnaires (ICIQ),[14] and International Continence Society (ICS) male and female questionnaires.[4,11] The majority of the content of such questionnaires is disease-specific

Accepted for publication June 2015

Acknowledgements
Drs. Elise De, Matthew Rutman and J. Christian Winters served as expert panel members and assisted with development of the questionnaire.

Address correspondence to Dr. Matthew S. Benedon, 445 East 77th Street, New York, NY 10075 USA

and, by definition, fails to adequately capture the entire range of a patient's symptoms. While questionnaires designed to broadly assess LUTS are available, most are quite extensive and may be impractical for administration. The aim of this study is to devise and assess a new LUTS questionnaire that broadly assesses LUTS in a concise fashion in terms of its validity, reliability and responsiveness to change.

Materials and methods

This institutional review board-approved validation study was performed at both an office-based practice and a university hospital outpatient department. An expert panel comprised of the two senior authors and three urologists expert in lower urinary tract dysfunction was convened to create a user-friendly, self-administered questionnaire that assesses the full complement of storage and voiding LUTS.

The topic exploration portion of content validation began with the panel members reviewing a list of LUTS descriptors generated from one-on-one patient interviews, which were then compared with a bank of questions and descriptors of bladder sensations retrieved from studies of three prior LUTS questionnaires conducted by the authors.[13,15,16] In addition, questions were obtained from other published studies of validated questionnaires and descriptors.[3,17] The first iteration was comprised of 23 items measuring both symptoms and frequency of symptoms chosen and modified by the expert panel from the questionnaires cited above. They were selected based on the current ICS definition of LUTS[18] and the most commonly noted symptoms reported during patient interviews. Question responses were designed on a 5-point Likert scale for symptom severity gradation. Expert opinion and subject feedback were incorporated via the following sequential method.

In all rounds of testing, subjects were restricted to adult men and women 18 years of age or older. Exclusion criteria included illiterate and non-English speaking patients, age < 18 years, and those deemed unreliable by the research staff. The first written draft of the questionnaire was administered on site to 30 consecutive subjects who presented with LUTS. Once each subject completed recording their answers on their paper questionnaire, they were immediately interviewed individually by members of the research staff regarding question clarity, simplicity and comprehensiveness, as well as the patient's opinion regarding the inclusion/exclusion of items in the questionnaire. They were specifically asked whether they had any symptoms that were not queried and, if so, asked to describe those symptoms in their own words.

Using such feedback, symptoms that were not covered by the existing questions were reviewed and the panel queried about the propriety of devising additional questions that addressed those symptoms. Additionally, the original 23 questions were revised accordingly by members of the expert panel and presented to each other for further review of clarity, content relevance and comprehensive coverage of all aspects of LUTS. All members of the expert panel provided their written responses to the subject-assisted revisions. The questionnaire items were then further edited based on comments from both the expert panel and study subjects and reduced to 14 items.

To assess test-retest, discriminant validity and its ability to detect change, the questionnaire was again administered to consecutive subjects recruited from the existing patient populations of the two outpatient facilities. Additional normal volunteers were recruited from employees, family members, urology residents and medical and college students.

The subjects were divided into three groups based on the medical records of patients seen by urologists who are expert in evaluating LUTS: 1) individuals without LUTS (normals), 2) patients with LUTS, but without urgency or urge incontinence, and 3) patients with OAB. LUTS patients with urgency and/or urge incontinence were classified as overactive bladder (OAB) in our cohort. The groups were created taking into consideration that OAB is a subcategory of LUTS. LUTS was defined as any combination of the following symptoms - urinary frequency, urgency, urge incontinence, nocturia, stress incontinence, difficulty voiding, hesitancy, weak stream, intermittency, and a feeling of incomplete bladder emptying. OAB was defined as urinary urgency with or without urge incontinence, urinary frequency and nocturia. Subjects completed the written questionnaire in privacy on two separate occasions during a period of 2-4 weeks (test #1 and test #2). During this time, no medical/surgical treatment intervention or invasive office procedure (i.e. cystoscopy, urodynamics) took place. Patients whose symptoms changed during this period were excluded.

To determine the questionnaire's responsiveness to change, a secondary study was conducted in which the LUTSS and a previously validated outcome measure, the Patient Global Impression of Improvement (PGI-I), were used as outcome measures for prostate surgery. This was a prospective, open observational study of a consecutive series of men who underwent transurethral resection (TURP) or KTP laser ablation (KTPLAP) for prostatic obstruction. Prostatic obstruction was defined by urodynamic criteria (Bladder Outlet Obstruction Index > 40) or BOOI < 40

for flow < 12 mL/s, a sustained detrusor contraction, and radiographic evidence of prostatic obstruction. The primary outcome measure was a change in the PGI-I. Secondary outcome measures included the total LUTSS, the OABSS, maximum urinary flow rate (Qmax) and post-void residual (PVR).

Statistics

Test-retest reliability was assessed separately for men, women, and the total cohort using intraclass correlation (ICC) based on a model with 2-way random effects and absolute agreement. Inter-item reliability was assessed using Cronbach's α and was calculated for the total score for men, women, and the total cohort. The standard deviation (SD) and mean for each question was reviewed to assess if any item exhibited ceiling or floor effects. A principal factor analysis was performed to detect subscales within the questionnaire. Discriminant validity was assessed by comparison of mean initial visit scores among the three groups using ANOVA with bootstrap-generated adjusted p values to deal with multiple post hoc tests. Effect size was calculated as the difference between mean scores for normal individuals versus LUTS patients and normal individuals versus OAB patients, divided by pooled standard deviation. Demographic characteristics of the study sample, such as age and gender, were compared among the three groups using Pearson chi-square test and Student's t-test. All statistical procedures were performed with $p < 0.05$ considered a priori statistically significant.

Pearson's correlation coefficient was used to assess criterion validity by correlating the storage and voiding subscales of the LUTSS with the corresponding questions addressing storage (questions 2, 4 and 7) and voiding symptoms (questions 1, 3, 5 and 6) in the AUASS, and also correlating the total score from both instruments. Consecutive patients completed these questionnaires on the same date.

In the responsiveness to change secondary study, a paired student t-test was used to compare preoperative and postoperative mean values.

Results

Content validation

Topic exploration

Thirty patients with LUTS completed the first version of the questionnaire. The group was comprised of 17 females and 13 males, mean age 56 years old (range 19-78), 18 Caucasians (60%), 6 African Americans (20%), 4 Hispanics (13%) and 2 Asians (7%), all of whom presented with LUTS. Seventy-three percent were college graduates, 20% graduated high school and 7% had lesser educational background. Twenty-five subjects stated the questions were clear, easy to understand, and addressed all of their symptoms. Five subjects claimed that the questionnaire did not address all of their symptoms. One patient did not understand the meaning of "bladder control". One question assessing intermittency was misinterpreted by two patients as their ability to voluntarily interrupt the urinary stream during voiding. One 76-year-old male subject did not understand the concept of urgency. Four subjects complained of a persistent "awareness" of the need to void, "always need to know where a bathroom is," "need to go before I go into a meeting or a car," and that these sensations were different than urgency. That "awareness" was not completely relieved by voiding. Two of these subjects also noted that shortly after voiding they sometimes had to "go again." Two suggested including the concept of post-void dribbling in the questionnaire.

The panel reviewed the comments and revised the questionnaire accordingly. While the importance of post-void dribbling and the persistent awareness of the need to void was recognized, neither was included in the final version of the questionnaire because, based on our current knowledge base, neither has a well-defined differential diagnosis nor effective treatment. Further, there was no way to fit "awareness" into the likert scoring system. The four questions relating to stress urinary incontinence (SUI) were combined into one question in order to eliminate some of the burden of the questionnaire. Three questions relating to daily frequency of specific symptoms were deleted for two reasons: 1) They were skip questions that presented excessive burden, possible confusion and 2) the clinical relevance of those questions was not significant enough to warrant the additional burden on the patient if included.

Content validation

The final LUTSS questionnaire consisted of 14 questions on a 5-point scale scored from 0-4. Nine questions pertained to storage symptoms, 4 to voiding symptoms and 1 to bother. The total score ranged from 0 to 56, where higher scores indicated worse symptoms.

Other validation

A total of 204 consecutive patients were recruited from both an office-based practice and a university hospital outpatient department. Eleven were excluded for not completing the final version of the questionnaire at visits 1 and 2, and 2 more due to urinary tract infections occurring between both administrations. Of the 191 remaining, 80 (42%) were males and 111 (58%) were females with a mean age of 65 years (range 22-91). The

study population consisted of 105 (55%) Caucasians, 57 (30%) African Americans, 19 (10%) Hispanics, and 10 (5%) Asians. Of the subjects, 72 (38%) were classified as OAB with or without other LUTS, 91 (48%) as LUTS without urgency, and 28 (14%) in the control group. There were significantly more women than men in the OAB group (71% versus 29%, $p < 0.05$) with a more equal gender distribution in the LUTS without urgency and normal groups (49% versus 51% and 57% versus 43%, $p > 0.05$). There was a significant difference in age between the control and OAB group, with the OAB group being older than the control (mean age 68 versus 59, $p < 0.05$).

Test-retest validation

A high level of test-retest reliability was shown, Table 1 between the two administrations of the questionnaire: intraclass correlation was 0.97 for men, 0.95 for women, and 0.96 for all patients. Strong internal consistency for the questionnaire was demonstrated with Cronbach's α of 0.79 in men only, 0.76 in women only, and 0.77 for the total cohort. A review of the mean and standard deviation for each item in the questionnaire (data not shown) did not show any items to exhibit a ceiling or floor effect where nearly all patients respond to an item in the same way. Standard deviations for each question ranged between 1 and 1.5, demonstrating reasonable variability.

Principal factor analysis revealed two primary factors that together account for 94% of variance shared among the 14 items in the questionnaire. These factors are orthogonal or mutually uncorrelated. Factor 1 accounted for 61% of variance and loaded on questions 3-6, 9, and 14. Cronbach's α for a subscale containing these items was 0.84 with a test-retest intraclass correlation of 0.97. Factor 2 accounted for 33% of variance and loaded on questions 10-13. Cronbach's α for a subscale containing these items was 0.74 with a test-retest intraclass correlation of 0.92. The remaining items, questions 1, 2, 7, and 8 had Cronbach's α 0.75 with a test-retest intraclass correlation of 0.96. Since questions 1-9 assess storage symptoms and questions 10-13 assess voiding symptoms, the results of the principal factor analysis suggest that the use of these particular subsets of questions as subscales can be considered appropriate and statistically valid. In addition, the storage subscale is subdivided into the OABSS (questions 1-6 and 9) and incontinence subscale (questions 6-8). These different subscales provide not only an overall symptom severity score, but also a metric to evaluate individual symptoms or a group of symptoms and may ultimately be used as a quantifiable outcome measure.

Discriminant validity

Discriminant validity was demonstrated by an ANOVA analysis and post-hoc bootstrap-generated adjustments. Significant differences ($p < 0.001$) in mean total scores were observed among the three groups, Table 1. Effect size, as measured by the ratio between the difference of mean scores for normal individuals versus LUTS patients and normal individuals versus OAB patients, divided by pooled standard deviation was 1.67 and 2.22, respectively.

Taking into account that OAB is a subcategory of LUTS, receiver operating characteristic analysis suggested that a LUTSS score of 15 and above, was able to discriminate between normal subjects and abnormal patients (including both OAB and LUTS groups) with an 84% sensitivity and 86% specificity.

Criterion validity

In order to evaluate criterion validity, 196 consecutive male patients, mean age 58 (range 36-79) were asked to complete the LUTSS and the AUASS on the same day. Pearson's correlation coefficient for the storage, voiding, and total score of each instrument were 0.46, 0.55 and 0.49 ($p < 0.01$) respectively; revealing a moderate, but statistically significant relationship across both the LUTSS and AUASS.

TABLE 1. **Test #1 and test #2 scores (n = 191)**

Group (n)	**Test #1‡ Mean score (SD)**	**Range**	**Test #2‡ Mean score (SD)**	**Range**
Normal (28)	8.6* (5.0)	0-20	9.0* (5.4)	0-21
LUTS (91)	19.9* (8.2)	3-38	19.8* (7.9)	3-38
OAB (72)	26.8* (8.8)	8-45	27.2* (8.9)	7-50

SD = standard deviation; LUTS = lower urinary tract symptoms; OAB = overactive bladder
‡No significant difference within group (test #1 versus test #2) based on non-pooled analysis
*$p < 0.001$ for omnibus test and for all post-hoc pair-wise tests (see text for details)

TABLE 2. **Pre and postoperative in men undergoing prostate surgery for prostatic obstruction (n = 84)**

	Mean preop score (SD)	Mean postop score (SD)
PGI-I	n/a	0.8
LUTSS (range 0-56)	21.5* (11.40)	11.3* (7.38)
OABSS (range 0-28)	11* (6.32)	6.6* (4.70)
Qmax (mL/sec)	6.6* (4.37)	15.2* (8.85)
PVR (mL)	387.1* (442.5)	43.9* (109.1)

SD = standard deviation; n/a = not applicable; PGI-I = patient global impression of improvement; LUTSS = lower urinary tract symptom score; OABSS = overactive bladder symptom score; PVR = post-void residual
*p< 0.01

Responsiveness to change

In our secondary study to determine the questionnaire's responsiveness to change, Table 2, a total of 84 men, mean age 64 (range 53-76) who underwent prostate surgery were evaluated. A successful outcome, defined as a PGI-I score of 0 to 2 was achieved in 81/84 (96% percent) of patients. The mean subjective assessment of improvement was 0.8 indicating that the average patient felt he was either "cured" (score of "0") or "very much improved" (score of "1"). All other secondary outcome measures improved accordingly. The decrease from pre to postoperative LUTSS mirrors the improvement demonstrated by the PGI-I, OABSS and uroflow parameters, demonstrating that the LUTSS was responsive to change.

Discussion

Why another questionnaire? Numerous questionnaires to evaluate LUTS have been devised and validated and the ICI has recommended that no further questionnaires be developed.[2] In our judgment , the ICI position is shortsighted and stifles innovation. We believe that all existing questionnaires have sufficient deficiencies to warrant consideration of another one.[3-12] The most important deficiency is that there is not a single existing questionnaire (except the OABSS[13] which is a precursor to the LUTSS) that is nuanced enough with respect to urinary urgency. All other questionnaires consider urinary urgency to be a pathologic symptom that is an "all or none phenomena," and that it cannot be graded. Yet there is ample evidence that intensity of the "desire to void" can be graded. Whether this graded desire to void rises to the level of urgency is a semantic, not a scientific question.[15,19,20] In our judgment, this granular approach to urgency and the desire to void is an integral part of understanding LUTS and should be an integral part of a LUTS questionnaire, which is exactly what the LUTSS provides. No other questionnaire takes this graded concept into account, describes how often it occurs, or its severity (i.e. how long the patient can postpone urination once the urge is experienced).

Another problem is that many questionnaires conflate symptoms with bother. For example, the ICIQ-UI,[5] one of the most commonly used questionnaires, has a maximum score of 21 for incontinence – 11 for the symptom and 10 for the bother. So, a patient can be bothered a lot by incontinence and score an 11/21 yet leak less often than once a week, or leak "all of the time" and not be bothered at all for a score of 10/21. Further, some questionnaires do not even ask about certain symptoms. The ICIQ-FLUTS does not ask about the strength of the urinary stream[11] and ICIQ-MLUTS does not ask about how often incontinence occurs.[4] Further, the ICIQ LUTS questionnaires ask slightly different questions for males and females. In The ICIQ-FLUTS, it asks "do you to strain to urinate, while the ICIQ-MLUTS asks "do you have to strain to continue to urinating." For men the question about stress incontinence is "does urine leak when you cough or sneeze and for women "does urine leak when you are physically, active, exert yourself or cough or sneeze."

The LUTSS presented herein is a 14-item questionnaire with a five-point Likert scale that has been developed by the authors over the course of over 3 years. Unlike many of the existing questionnaires, it has been validated in both men and women. Further, it pertains to the entire spectrum of common LUTS, excluding, for reasons discussed above, post-void dribbling, enuresis and a constant, uncomfortable awareness of the need to void. There are four questions pertaining to voiding symptoms and nine relating to storage symptoms. Unique to the LUTSS are the seven questions that comprise the overactive bladder symptom score (part of the storage symptom questions). Voiding symptoms are not as nuanced as storage symptoms, therefore there are only

four questions pertaining to them, similar in content to other questionnaires.

Another unique feature of the LUTSS is its ability to aid in diagnosis. For example, the overactive bladder symptom score has already been separately validated to correlate with OAB severity - a score of 8 or more had a high correlation with the OAB symptom complex.[13] While the other subscales (storage, incontinence – and voiding) have not yet been validated, we are hopeful that they will be useful metrics for the underlying symptoms that ultimately can be used as a PRO tool. Furthermore, OAB and incontinence may be scored separately and/or grouped within the storage subscale following the ICS definitions of storage symptoms.[18]

One of the most commonly used PROs is the AUASS[3] and the LUTSS had acceptable criterion validity with it. Nevertheless, we believe that the AUASS storage questions are neither intuitive nor descriptive enough and do not provide a useful metric to quantify OAB symptoms. For example, it asks the questions: "How often have you had to urinate less than every two hours?" And, "how often have you found it difficult to postpone urination?" This does little to discern how frequently and why they void, whether or not they experience urgency and, if so, how severe it is.

The fifth International Consultation on Incontinence gave a grade A recommendation to its own proprietary ICIQ modular questionnaires.[2] They describe 14 questionnaires currently in use, with separate ones for men and women (ICIQ-MLUTS and ICIQ-FLUTS).[4,11] These questionnaires are meant to be used together to evaluate LUTS. So, in contrast to a single questionnaire that encompasses the entire scope of LUTS (the LUTSS), the clinician or researcher has to pick from a large menu of questionnaires that may prove burdensome for clinical practice. Further, the ICIQ utilizes a qualitative scale (never, occasionally, sometimes, most of the time and all the time) that may be interpreted differently among patients; whereas, the LUTSS quantifies events with more precision, e.g. never, a few times a month, a few times a week, at least once a day. So, for example, two patients who are incontinent a few times a week, may answer the ICIQ occasionally or sometimes; we believe that this ambiguity makes the LUTSS a more robust and accurate PRO. Also, the LUTSS is valid for both genders without the need to use one for males and one for females. Further, the ICIQ-MLUTS and ICIQ-FLUTS include 13 and 12 bother questions each, making them 26 and 24 item questionnaires, respectively, not 12 and 13 item questionnaires as advertised. We believe, despite their rigorous psychometric validation, these questionnaires fail to provide a nuanced response when asking about urgency and at the same time does not discern its cause and severity. Furthermore, the question: "Does urine leak before you can get to the toilet?" is not specific for any type of incontinence; and even though they include a voiding and incontinence sub-score, they do not provide a sub-score to quantify the spectrum of storage symptoms.

The LUTSS asks a single question about bother that relates to the entire LUTS experience and places greater emphasis on symptoms rather than bother, and omits quality of life (QoL) altogether. There are several reasons for this. While QoL and bother are important parameters in their own right, an abundance of these questions may dilute efficacy outcomes when combined in a single symptom score, and might be considered burdensome by some individuals and might lead to missing data. Further, there is an unpredictable relationship between objective measures of symptom severity, bother and QoL, such that two subjects with identical symptoms and severity would have very different symptom scores depending on their degree of bother and effect on QoL.[13,21] But, neither bother nor QoL have anything to do with the severity of symptoms or the underlying conditions. While we agree that these domains are very important and should be assessed, neither are direct nor objective measures of symptom severity, and their inclusion dilutes the impact of the severity score. That is why a single global patient bother question was deemed worthy of inclusion, in accordance with ICI recommendations,[2] while minimizing burden for patients.

The LUTSS had excellent test-retest reliability for both individual items as demonstrated by the high intraclass correlation coefficients for men, women, and the total cohort, meaning stable responses are expected as long as no intervention takes place. No ceiling or floor effect was detected for any item; therefore we can be confident that all items will elicit a broad range of responses from the population.

The LUTSS demonstrated high discriminant validity between normals, LUTS, and OAB patients. When comparing these groups to each other, there were significant differences in the scores. This was also demonstrated by measured effect size, as an estimate of the difference between two population's mean score. It showed patients with LUTS had mean scores 1.67 SD above normal individuals and OAB patients had mean scores 2.22 SD above normals. However, this discrimination between normal individuals and abnormal patients based only on questionnaires should not be considered a binary phenomenon. Some patients may not report symptoms during history taking, but have a LUTSS above our cut off of 14 and vice versa. This can be explained in part by the questionnaire's sensitivity and specificity (84% and 86%, respectively),

the patient's awareness of symptoms compared to objective measures (i.e. Qmax, PVR, urodynamics, etc.) and the degree of bother.

Regarding its ability to detect change, the scores obtained in the LUTSS in men after undergoing prostate surgery were significantly lower than preoperative scores, Table 2, and improvement was validated by other metrics (PGI-I, OABSS, Qmax and PVR).

There are a number of shortcomings of the current study. There were differences in demographic characteristics among the three groups as outlined in the results section. Despite its responsiveness to change in men undergoing prostate surgery, further studies should be conducted to assess the questionnaire's responsiveness in female specific diseases and disease specific treatments (i.e. alpha blockers, anticholinergics, bladder training, slings, etc.). In addition, further studies are needed to define relevant cut off values for each of the symptom subscales.

Conclusion

The LUTSS is a validated questionnaire that assesses a full range of LUTS symptoms. Unlike other published questionnaires, to our knowledge, it is the only one that takes a granular and nuanced approach to overactive bladder. The LUTSS does not consider urinary urgency to be an all or none phenomena, but rather, a symptom that can be graded. Content validity has been established. It has excellent criterion and test-retest validity, internal consistency and is responsive to change. The LUTSS has already proven to be a valid PRO for prostate surgery, is a concise and easily administered tool, and therefore practical for use in clinical practice and research purposes.

Disclosures

JGB: Consultant for Astellas and Elsevier; Shareholder in P Square Medical; Shareholder and co-owner of intellectual property with Symptelligence; provided expert testimony in mesh litigation lawsuits
JMW: Shareholder in Symptelligence
JPW: Consultant for Pfizer, Ferring, Astellas, Allergan, Vantia, Symptelligence, Elsevier
JFT, GM, MSB, BL, FMF, JW: Nothing to disclose □

References

1. Gordon D, Groutz A. Evaluation of female lower urinary tract symptoms: overview and update. *Curr Opin Obstet Gynecol* 2001; 13(5):521-527.
2. Kelleher C, Staskin DR, Cherian P et al. Patient Reported Outcome Assessment. In: Abrams P, Cardozo L, Khoury S, Wein A, editors. Incontinence. 5th ed: EAU; 2013. p. 389-429.
3. Barry MJ, Fowler FJ Jr, O'Leary MP et al. The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. *J Urol* 1992;148(5):1549-57; discussion 64.
4. Donovan JL, Abrams P, Peters TJ et al. The ICS-'BPH' Study: the psychometric validity and reliability of the ICSmale questionnaire. *Br J Urol* 1996;77(4):554-562.
5. Avery K, Donovan J, Peters TJ, Shaw C, Gotoh M, Abrams P. ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. *Neurourol Urodyn* 2004;23(4):322-330.
6. Donovan JL, Peters TJ, Abrams P, Brookes ST, de la Rosette JJ, Schafer W. Scoring the short form ICSmaleSF questionnaire. International Continence Society. *J Urol* 2000;164(6):1948-1955.
7. Coyne K, Revicki D, Hunt T et al. Psychometric validation of an overactive bladder symptom and health-related quality of life questionnaire: the OAB-q. *Qual Life Res* 2002;11(6):563-574.
8. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fantl JA. Health-related quality of life measures for women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Continence Program in Women (CPW) Research Group. *Qual Life Res* 1994;3(5):291-306.
9. Uebersax JS, Wyman JF, Shumaker SA, McClish DK, Fantl JA. Short forms to assess life quality and symptom distress for urinary incontinence in women: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Continence Program for Women Research Group. *Neurourol Urodyn* 1995;14(2):131-139.
10. Sandvik H, Hunskaar S, Seim A, Hermstad R, Vanvik A, Bratt H. Validation of a severity index in female urinary incontinence and its implementation in an epidemiological survey. *J Epidemiol Community Health* 1993;47(6):497-499.
11. Jackson S, Donovan J, Brookes S, Eckford S, Swithinbank L, Abrams P. The Bristol Female Lower Urinary Tract Symptoms questionnaire: development and psychometric testing. *Br J Urol* 1996;77(6):805-812.
12. Hald T, Nordling J, Andersen JT, Bilde T, Meyhoff HH, Walter S. A patient weighted symptom score system in the evaluation of uncomplicated benign prostatic hyperplasia. *Scand J Urol Nephrol Suppl* 1991;138:59-62.
13. Blaivas JG, Panagopoulos G, Weiss JP, Somaroo C. Validation of the overactive bladder symptom score. *J Urol* 2007;178(2):543-547; discussion 7.
14. Abrams P, Avery K, Gardener N, Donovan J, Board IA. The International Consultation on Incontinence Modular Questionnaire: www.iciq.net. *J Urol* 2006;175(3 Pt 1):1063-1066;discussion 6.
15. Blaivas JG, Panagopoulos G, Weiss JP, Somaroo C, Chaikin DC. The urgency perception score: validation and test-retest. *J Urol* 2007;177(1):199-202.
16. Weiss JP, Blaivas JG, Tash Anger JA, Di Blasio CJ, Panagopoulos G, Gerboc J. Development and validation of a new treatment outcome score for men with LUTS. *Neurourol Urodyn* 2004;23(2):88-93.
17. Das R, Buckley JD, Williams MT. Descriptors of sensation confirm the multidimensional nature of desire to void. *Neurourol Urodyn* 2015;34(2):161-166.
18. Abrams P, Cardozo L, Fall M et al. The standardisation of terminology of lower urinary tract function: report from the Standardisation Sub-committee of the International Continence Society. *Neurourol Urodyn* 2002;21(2):167-178.
19. Cardozo L, Coyne KS, Versi E. Validation of the urgency perception scale. *BJU Int* 2005;95(4):591-596.
20. Coyne KS, Margolis MK, Hsieh R, Vats V, Chapple CR. Validation of the urinary sensation scale (USS). *Neurourol Urodyn* 2011;30(3):360-365.
21. Scarpero HM, Fiske J, Xue X, Nitti VW. American Urological Association Symptom Index for lower urinary tract symptoms in women: correlation with degree of bother and impact on quality of life. *Urology* 2003;61(6):1118-1122.