
Validation of the Overactive Bladder Symptom Score

Jerry G. Blaivas,* Georgia Panagopoulos, Jeffrey P. Weiss† and Chandra Somaroo

From the Weill Medical College of Cornell University (JGB, JPW), Lenox Hill Hospital (JGB, GP, JPW) and UroCenter of New York (JGB, JPW, CS), New York, New York

Purpose: We validated a new 7-item overactive bladder symptom score.

Materials and Methods: Normal subjects and consecutive patients with lower urinary tract symptoms with or without overactive bladder were recruited and classified into 3 groups based on their response to the question on an intake questionnaire, "Do you ever experience a sudden urge to urinate that makes you want to stop what you are doing and rush to a bathroom?" Subjects completed the written questionnaire in privacy on 2 occasions.

Results: There were 84 subjects, including 33 men and 51 women, with a median age of 62 years (range 18 to 88). Of the subjects 33 (39%) had overactive bladder, 30 (36%) had lower urinary tract symptoms without overactive bladder and 21 (25%) were normal. There was a high level of internal consistency at visits 1 and 2 (Cronbach's $\alpha = 0.83$ and 0.80 , respectively, $p < 0.001$). For test-retest reliability Spearman's rank order correlation coefficients for the items were $r = 0.72$ to 0.79 ($p < 0.001$). A strong correlation was also observed between the total 7-item scores at visits 1 and 2 for each diagnostic subgroup, and for all participants ($r = 0.86$, $p < 0.001$). Discriminant validity was established by determining significant differences in responses among the 3 subgroups at each administration ($p < 0.01$).

Conclusions: The overactive bladder symptom score is a valid instrument that assesses all aspects of overactive bladder. It may be used as a symptom score.

Key Words: bladder, symptoms, questionnaire, urination disorders, urinary incontinence

As defined by the International Continence Society, OAB is "urgency, with or without urge incontinence, usually with frequency and nocturia."¹ Although there are many validated lower urinary tract symptom questionnaires, of which some are devoted to OAB and/or incontinence, there are no validated OAB symptom scores that quantitate all aspects of OAB and none that include a graded response for urgency. We report the validation of a questionnaire and symptom score that quantitates all aspects of OAB and includes a graded response for urgency.

MATERIALS AND METHODS

This institutional review board approved validation study was performed at 2 clinical urology sites in the Northeastern United States. Inclusion criteria were 1) normal volunteers recruited from medical staff and family members of subjects, 2) patients with LUTS without OAB, and 3) subjects with LUTS and OAB. The OABSS questionnaire was developed for use among ethnically diverse English speaking men and women 18 years or older who can read at a fourth grade level.

An expert panel was assembled to select and develop the questions that comprised the OABSS. Content validity was

established by performing 5 steps that incorporated expert opinion and subject feedback.

1) The first iteration comprised 6 questions written by 2 of us (JB and CS) based on the International Continence Society definition of overactive bladder, other LUTS questionnaires and 1 question previously developed for the urge perception score.^{2,3}

2) Approximately 20 consecutive subjects who presented with urgency symptoms were administered the written questionnaire on site and were interviewed by 1 of 2 members of the research staff (CS or JB) immediately afterward. The researchers reviewed subject responses with the subject and asked about the clarity of each question (yes/no), whether it could be phrased more simply and whether the subject had experienced any additional symptoms that were not included in the questionnaire. When discrepancies were observed between subject responses to this questionnaire, the intake questionnaire and the clinical impression of the examining physician, the definition of urgency was presented to the subject and further discussion ensued.

3) Using such feedback from subjects the questions were revised by JB and CS, and presented to each panel member for review of clarity, content relevance and comprehensive coverage of all aspects of OAB.

4) Panel members provided their written responses to the subject assisted revision. These were evaluated by 2 researchers (JB and CS), who further edited the questions based on these comments.

5) The entire panel subsequently convened and reviewed the revised questions in detail. Disagreements were resolved by discussion among the entire panel until unanimous approval was obtained. We did not assess agreement among

Submitted for publication January 11, 2007.

Study received institutional review board approval.

Supported by the Institute for Bladder and Prostate Research, Novartis and Pfizer.

* Financial interest and/or other relationship with Novartis, Pfizer, Johnson & Johnson, and Endogun.

† Financial interest and/or other relationship with Pfizer.

the panel by calculating the content validity ratio because we required unanimous agreement.

In its final form OABSS was developed as a self-administered questionnaire consisting of 7 questions on a 5-point Likert scale. Five questions are related to urinary urgency and 2 are related to daytime and nighttime urinary frequency (see Appendix).

All subjects were interviewed by a research associate (CS) to ensure that they met inclusion criteria for their respective groups. Since OAB is a subcategory of LUTS, subjects with LUTS were subdivided into 2 groups, including those with LUTS without urgency and those with OAB with or without other LUTS. LUTS consists of urinary frequency, urgency, urge incontinence, dysuria, hesitancy, weak stream, intermittence and a feeling of incomplete bladder emptying. To be included in the LUTS without urgency group a subject must have had any or all of the LUTS symptoms listed except urgency or urge incontinence. If the subject had urgency or urge incontinence, that subject was included in the OAB group whether or not there were other LUTS.

Subjects were assigned to the appropriate group based on the clinical diagnosis, which was obtained as follows. Subjects completed an intake LUTS questionnaire and a validated OAB and incontinence questionnaire, and underwent direct questioning by a research associate. Each of these questionnaires has at least 1 question that paraphrases the International Continence Society definition of urgency, for example "Do you ever experience a sudden urge to urinate that makes you want to stop what you are doing and rush to a bathroom?" If there was complete agreement between each of the questionnaires and the subject history, the subject was classified accordingly. If there was not complete agreement, the subject was questioned again, the symptom urgency was clarified and the subject was reclassified based on the verbal response.

Subjects completed the written questionnaire in privacy on 2 occasions during 3 to 10 days during which there was no change in symptoms. Subjects were excluded if there was any reason to believe that symptoms would not be stable, if the subject stated that symptoms had changed between the first and second administration of the questionnaire or if they failed to answer all questions.

Statistical Analysis

Internal consistency, based on the average interitem correlation, was calculated using Cronbach's α statistic (intra-class correlation coefficient). Test-retest reliability was determined by examining the association of responses to each question at visits 1 and 2 using Spearman's rank order correlation coefficient because the 7 OABSS items were measured in an ordinal (rank order) rather than a continuous scale.

Discriminant validity was assessed by calculating the average total scale score for each of the 3 groups and comparing these scores using 1-way ANOVA and Fisher's LSD post hoc test. Demographic characteristics of the study sample were compared using ANOVA for continuous data and the Pearson chi-square test for categorical data. All statistical procedures were performed using SPSS®, version 14.0 with $p < 0.05$ considered a priori statistically significant.

RESULTS

Study Sample

A total of 90 consecutive participants were recruited and completed visits 1 and 2. Six subjects were excluded because they failed to answer 1 or more questions. Of the remaining 84 subjects 33 were men (39%) and 51 were women (61%) with a median age 62 years (range 18 to 88). Of the subjects 33 (39%) had OAB, 30 (36%) had LUTS without OAB and 21 (25%) were normal. There was a significant difference in gender and age distributions across the 3 groups. There were more women than men in the OAB group, whereas men and women were approximately equally distributed in the LUTS and normal groups (88% vs 12% and 57% vs 43%, respectively, $p < 0.001$). The OAB and LUTS groups were significantly older than the normal group (median age 72, range 46 to 88 and 71, range 29 to 81, respectively, vs 33, range 18 to 70, $p < 0.001$). There was no difference in the median age between the OAB and LUTS groups ($p > 0.10$).

Psychometric Properties

A high level of internal consistency was observed among the 7 OABSS items in the total cohort at visits 1 and 2 (Cronbach's $\alpha = 0.83$ and 0.80 , respectively, $p < 0.001$). Table 1 shows the correlation coefficients for each of the 7 question items at visit 1.

Test-Retest Reliability

The observed Spearman's rank order correlation coefficient indicated a strong association between participant responses to each of the 7 items at visits 1 and 2, respectively. Spearman's rank order correlation coefficients were $r = 0.72$ to 0.79 ($p < 0.001$). A strong correlation was also observed between the total 7-item score at visits 1 and 2 for each diagnostic subgroup and for all participants ($r = 0.86$, $p < 0.001$). Table 2 lists these results.

Comparison of the average total scores obtained for all participants at visits 1 and 2 did not show statistical significance (10.4 ± 5.1 vs 10.4 ± 5.3 , $p = 1.0$). This provided additional support that participant responses did not differ between the 2 visits.

Discriminant Validity

Discriminant validity was assessed by comparing average total scores among the 3 diagnostic subgroups at visits 1 and 2 using ANOVA statistical methodology. Analysis indicated that there were significant differences in the responses among the 3 subgroups at the 2 assessment periods ($p < 0.01$). Table 3 shows the mean \pm SD of these analyses.

TABLE 1. Nine question inter-item correlation matrix at visit 1

Question No.	Correlation (question No.)						
	1	2	3	4	5	6	7
1	1.000						
2	0.18	1.000					
3	0.35*	0.39*	1.000				
4	0.37*	0.39*	0.63*	1.000			
5	0.34*	0.45*	0.51*	0.58*	1.000		
6	0.01	0.40*	0.29*	0.28†	0.50*	1.000	
7	0.20	0.29*	0.39*	0.39*	0.59*	0.49*	1.000

* Significant at the 0.01 level.

† Significant at the 0.05 level.

TABLE 2. Spearman's rank order correlation coefficients of agreement between scores at visits 1 and 2

OABSS Question No.	Spearman's Coefficient
1	0.72
2	0.73
3	0.74
4	0.72
5	0.73
6	0.77
7	0.79
Total score:	0.86
LUTS	0.84
OAB	0.71
Normal	0.80

All values significant at $p < 0.001$.

DISCUSSION

The third ICI recommended that symptom questionnaires should include domains relating to symptoms, quality of life and bother.⁴ We agree in principle but believe that first and foremost a symptom questionnaire should address symptoms. Furthermore, although there are a number of validated questionnaires that deal with LUTS, none were developed for use as a symptom score and none combine a graded response for urgency with all of the symptoms of OAB.⁵⁻¹³ Although some investigators believe that urgency is an all or none phenomenon that cannot be graded,¹⁴ recent studies showed that urgency is a subjective symptom that can be graded.^{3,15-17}

As reported, the OABSS includes not only a graded response for urgency, but also all of the symptoms of OAB, and it was specifically designed for use as a symptom score. The 7 questions relate to all symptoms of OAB, including 1 each on urinary frequency and nocturia, 3 on urgency, 1 on urge incontinence and 1 generic question about bladder control. The total score range is 0 to 28. The higher the score, the worse the symptoms. In addition to the complete OABSS, there is an urgency subscale (questions 3 to 6) that can be used to grade the severity of urgency.

TABLE 3. Subgroup total scores at visits 1 and 2

Diagnostic Group	Mean \pm SD Visit 1*	Mean \pm SD Visit 2†
Normal	5.6 \pm 2.8	5.1 \pm 2.5
LUTS	10.1 \pm 4.2	11.0 \pm 4.4
OAB	13.5 \pm 4.9	13.2 \pm 4.5

* Significant at $p < 0.001$.

† LUTS vs. OAB $p < 0.01$ at visits 1 and 1, normal vs. OAB and LUTS $p < 0.001$ at visits 1 and 2, and LUTS vs. OAB $p < 0.05$ at visit 2.

We chose to omit quality of life and bother from the questionnaire for 2 reasons. 1) A number of validated instruments have been devised that already do that well.^{5-13,18,19} 2) We believe that, while they are important in their own right, QOL and bother domains tend to dilute efficacy parameters when combined in a single symptom score. For this reason we prefer to administer them as separate outcome instruments. For example, the short form of OAB-Q⁷ asks, "During the past 4 weeks, how bothered were you by" 6 OAB symptoms. There can be a poor relationship between symptom severity and objective measures, bother and QOL²⁰ so that, if a subject has severe incontinence but is not bothered by it, the total symptom score would be lower than if the subject was bothered by it if QOL was included. Thus, 2 subjects with exactly the same frequency and amount of incontinence could have different incontinence outcome scores if bother was used as a criterion.

The ICI further recommended that questionnaires should be "valid, reliable and responsive to change following standard psychometric testing."⁶ In this study OABSS was shown to be valid and reliable. Furthermore, it had excellent discriminate validity. There was clear separation in OABSS scores between normal subjects,⁶ and those with LUTS without OAB¹² and OAB¹⁷ (each $p < 0.001$). The responsiveness of OABSS to change has not yet been determined but it is currently under investigation in an ongoing study.

There are several shortcomings to the current study. 1) The demographic characteristics of the 3 groups were different with respect to age and gender, which was a natural outcome, given the demographics of the pathological conditions studied. 2) As alluded to, the responsiveness to change (treatment) was not studied.

Based on a large literature review only 2 incontinence questionnaires were highly recommended by ICI for use in men and women, that is ICI-Q⁵ and OAB-q.⁷ However, as alluded to, the latter does not actually quantify symptoms, only bother. We hope and believe that in its current iteration OABSS would prove to be a more useful tool for quantifying symptoms and outcome after treatment in men and women with OAB. Its advantages over existing scores are several. 1) It quantifies all OAB symptoms. 2) It provides a more detailed evaluation of the urgency symptom than any of the other questionnaires. 3) Unlike the other scores, it permits a graded response (a severity score) for urgency rather than the yes/no characterization offered by other instruments.

ACKNOWLEDGMENTS

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

APPENDIX

OAB Questionnaire

NAME: _____ DATE: _____

1. How often do you usually urinate during the day?

no more often than once in 4 hours

about every 3–4 hours

about every 2–3 hours

about every 1–2 hours

at least once an hour

2. How many times do you usually urinate at night (from the time you go to bed until the time you wake up for the day)?

0–1 times

2 times

3 times

4 times

5 or more times

3. What is the reason that you usually urinate?

out of convenience (no urge or desire)

because I have a mild urge or desire (but can delay urination for over an hour if I have to)

because I have a moderate urge or desire (but can delay urination for more than 10 but less than 60 minutes if I have to)

because I have a severe urge or desire (but can delay urination for less than 10 minutes if I have to)

because I have desperate urge or desire (must stop what I am doing and go immediately)

4. Once you get the urge or desire to urinate, how long can you usually postpone it comfortably?

more than 60 minutes

about 30–60 minutes

about 10–30 minutes

a few minutes (less than 10 minutes)

must go immediately

5. How often do you get a sudden urge or desire to urinate that makes you want to stop what you are doing and rush to the bathroom?

never

rarely

a few times a month

a few times a week

at least once a day

6. How often do you get a sudden urge or desire to urinate that makes you want to stop what you are doing and rush to the bathroom but you do not get there in time (ie you leak urine or wet pads)?

never

rarely

a few times a month

a few times a week

at least once a day

7. In your opinion how good is your bladder control?

perfect control

very good

good

poor

no control at all

Abbreviations and Acronyms

ICI	=	International Consultation on Incontinence
LUTS	=	lower urinary tract symptoms
OAB	=	overactive bladder
OABSS	=	OAB symptom score

REFERENCES

- Abrams P, Cardozo L, Fall M, Griffiths D, Rosier P, Ulmsten U et al: The standardisation of terminology of lower urinary tract function: report from the Standardisation Subcommittee of the International Continence Society. *Neurourol Urodyn* 2002; **21**: 167.
- Blaivas J, Panagopoulos G, Weiss JP, Somaroo C and Chaikin DC: The urgency perception score (UPS): validation and test-retest. *J Urol* 2007; **177**: 199.
- De Wachter S and Wyndaele JJ: Frequency-volume charts: a tool to evaluate bladder sensation. *Neurourol Urodyn* 2003; **22**: 638.
- Jonovan J, Bosch R, Gotoh M, Jackson S, Naughton M, Radley S et al: Incontinence, 2nd ed. Symptom and Quality of Life Assessment, Committee 6. Edited by P Abrams, L Cardozo, S Khoury and A Wein. Plymouth, United Kingdom: Health Publication Ltd 2005.
- Avery K, Donovan J, Peters TJ, Shaw C, Gotoh M and Abrams P: ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. *Neurourol Urodyn* 2004; **23**: 322.
- Donovan JL, Peters TJ, Abrams P, Brooks ST, de la Rosette JJ and Schafer W: Scoring the short form ICSmaleSF questionnaire. *International Continence Society. J Urol* 2000; **164**: 1948.
- Coyne K, Revicki D, Hunt T, Corey R, Stewart W, Bentkover J et al: Psychometric validation of an overactive bladder symptom and health related quality of life questionnaire; the OAB-q. *Qual Life Res* 2002; **11**: 563.
- Shumaker SA, Wyman JF, Uebersax JS, McClish D and 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**: 291.
- Uebersax JS, Wyman JF, Shumaker SA, McClish DK and 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**: 131.
- Sandvik H, Hunskaar S, Seim A, Hermstad R, Vanvik A and Bratt H: Validation of a severity index in female urinary incontinence and its implementation in an epidemiological survey. *J Epidemiol Community Health* 1993; **47**: 497.
- Jackson S, Donovan J, Brookes S, Eckford S, Swithbank L and Abrams P: The Bristol Female Lower Urinary Tract Symptoms questionnaire: development and psychometric testing. *Br J Urol* 1996; **77**: 805.
- Donovan JL, Abrams P, Peters TJ, Kay HE, Reynard J, Chapple C et al: The ICS-BPH Study: the psychometric validity and reliability of the ICSmale questionnaire. *Br J Urol* 1996; **77**: 554.
- Hald T, Nordling J, Andersen JT, Bilde T, Meyhoff HH and Walter S: A patient weighted symptom score system in the evaluation of uncomplicated benign prostatic hyperplasia. *Scand J Urol Nephrol, suppl.*, 1991; **138**: 59.
- Chapple CR, Artibani W, Cardozo LD, Castro-Diaz D, Craggs M, Haab F et al: The role of urinary urgency and its measurement in the overactive bladder symptom syndrome: current concepts and future prospects. *BJU Int* 2005; **95**: 335.
- Nixon A, Coleman S, Sabounjian L, Sandage B, Schwiderski UE, Staskin DR et al: A validated patient reported measure of urinary urgency severity in overactive bladder for use in clinical trials. *J Urol* 2005; **174**: 604.
- Oliver S, Fowler C, Mundy A and Craggs M: Measuring the sensations of urge and bladder filling during cystometry in urge incontinence and the effects of neuromodulation. *Neurourol Urodyn* 2003; **22**: 7.
- Zinner N, Harnett M, Sabounjian L, Sandage B Jr, Dmochowski R and Staskin D: The overactive bladder-symptom composite score: a composite symptom score of toilet voids, urgency severity and urge urinary incontinence in patients with overactive bladder. *J Urol* 2005; **173**: 1639.
- Kelleher CJ, Cardozo LD, Khullar V and Salvatore S: A new questionnaire to assess the quality of life of urinary incontinent women. *Br J Obstet Gynaecol* 1997; **104**: 1374.
- Stach-Lempinen B, Kirkinen P, Laippala P, Metsanoja R and Kujansuu E: Do objective urodynamic or clinical findings

determine impact of urinary incontinence or its treatment on quality of life? *Urology* 2004; **63**: 67.

20. van Brummen HJ, Heintz AP and van der Vaart CH: The association between overactive bladder symptoms and objective parameters from bladder diary and filling cystometry. *Neurourol Urodyn* 2004; **23**: 38.

EDITORIAL COMMENT

These authors propose a symptomatic analysis of patients with lower urinary tract symptoms (OAB symptoms). As such, this is a validated and stringently tested questionnaire that the authors meticulously developed to evaluate OAB symptomatology as a component of LUTS.

What is most important at this time is to realize that symptom analysis for LUTS, specifically OAB symptomatology, should not fall into the morass that has eventuated with quality of life questionnaires for LUTS. The problem is that so many questionnaires exist that there is no general agreement on which quality of life questionnaires are ideal. This makes cross-comparability among studies difficult and it also precludes the development of an ideal questionnaire that can be used generically.

With the recent focus on the importance of symptom analysis and appraisal of symptom change after intervention for LUTS, it is of paramount importance to realize the necessity of establishing a generally accepted symptom ap-

praisal tool. Although the currently derived tool appears to be inclusive, there are other tools in development or that were recently put forward that also attempt to accomplish this. Whether this tool is superior to those is problematic. Nonetheless, this tool appears to offer facility and more importantly global incorporation of symptom analysis across the OAB subcomponent complaints. It cannot be argued that appropriate symptom analysis is an important outcome measure for LUTS because these analyses represent what the individual experiences after therapy. Until we have general agreement about how best to measure the unique individual response to therapy we will continue to see the development of appraisal instruments such as this. These types of tools should be used in conjunction with 1-item symptom appraisals because the 1-item method provides the most rapid and clinically reflective responses. It is important that clinical researchers should remember that we are dealing with syndromes (a constellation of symptoms). Therefore, not only should an aggregate analysis be performed of those symptoms, but also a unique generic overview appraisal, such as is offered by 1-item appraisal questionnaires.

Roger Dmochowski

Department of Urologic Surgery
Vanderbilt University Medical Center
Nashville, Tennessee