Validation of the Urgency, Weak Stream, Incomplete Emptying, and Nocturia (UWIN) Score Compared With the American Urological Association Symptoms Score in Assessing Lower Urinary Tract Symptoms in the Clinical Setting

Khadijah Eid, Kevin Krughoff, Diliana Stoimenova, Daniel Smith, Jason Phillips, Colin O'Donnell, and Al Barqawi

OBJECTIVE To validate the Urgency, Weak stream, Incomplete emptying, and Nocturia (UWIN) survey for patients with lower urinary tract symptoms (LUTS) by comparison with the American Urological Association Symptoms Score (AUA-SS). The hypothesis is that the UWIN will perform as well as the AUA-SS in assessing LUTS symptoms and quality of life. The AUA-SS is complex for many patients and can be misunderstood. The UWIN questionnaire was developed to serve as a simpler and shorter version of the AUA-SS, with the intent of improving accuracy and minimizing error in assessing LUTS. The UWIN consists of 4 questions scored 0-3 to give a maximum score of 12.

METHODS We screened 700 patients in the urology clinic between 2011 and 2012. We enrolled 593 patients who completed the AUA-SS survey and UWIN in the same clinic visit. The AUA-SS and UWIN responses were evaluated using Spearman correlation coefficients and Bland-Altman graphs.

RESULTS Correlation coefficients were calculated between the corresponding AUA-SS and UWIN items on 593 matched surveys, demonstrating a strong correlation coefficient of 0.81 or greater for each question, which was statistically significant ($P < .0001$). The correlation coefficient between the total scores of the AUA and UWIN was 0.89 ($P < .01$). A second analysis was performed using Bland-Altman plots between AUA-SS and UWIN including total score, quality of life, and categories, which showed a good agreement.

CONCLUSION The UWIN appears to provide results comparable to the AUA-SS, while using a simpler format and taking less time to complete. UROLOGY 83: 181–185, 2014. © 2014 Elsevier Inc.

Benign prostate hyperplasia (BPH) is a common benign urologic condition in men, and its prevalence increases with age.[1] Deciding the correct course of management is imperative because BPH can profoundly affect a patient's quality of life (QOL), leading to further health complications and adding to increasing costs on the health care system. Clinical workup aids in deciding a proper course of treatment for patients such as watchful waiting or medical or surgical therapies.[2]

In the 2010 *Guidelines for the Management of BPH*, the American Urological Association (AUA) advocates the use of the Symptoms Score (AUA-SS) survey in addition to medical history, physical checkup, and investigation for the initial evaluation of the BPH patient. The AUA-SS gives clinicians the ability to objectively evaluate LUTS severity, response to therapy, follow-up duration, and symptom progression.[3] In addition to quality of life (QOL), the survey focuses on 7 areas: incomplete emptying, frequency, intermittency, urgency, weak urinary stream, hesitancy, and nocturia.[4]

Financial Disclosures: *The authors declare that they have no relevant financial interests.*

From the Division of Urology, Department of Surgery, University of Colorado Denver School of Medicine, Aurora, CO

Reprint requests: Al Barqawi M.D., F.R.C.S., Division of Urology, Department of Surgery, University of Colorado Denver School of Medicine, Academic Office One Bldg. Mail Stop C-319, 12631 E. 17th Ave., Rm. L15-5602 Aurora, CO 80045. E-mail: al.barqawi@ucdenver.edu

Submitted: June 28, 2013, accepted (with revisions): August 20, 2013

0090-4295/14/$36.00

http://dx.doi.org/10.1016/j.urology.2013.08.039

Patients provide a score ranging from 0-5 points for each question, which are summed to provide a total score of 0-35. As determined from the total score, patients fall into 1 of 3 categories of LUTS: mild (1-7), moderate (8-19), and severe (20-35). Although beneficial from a clinical standpoint, the AUA-SS has certain shortcomings. Perhaps the greatest shortcoming is the potential for patients to misunderstand and misinterpret the questions due to their length and wording, thereby resulting in inaccurate management and inappropriate care.

A shorter questionnaire, called the UWIN (Urgency, Weak stream, Incomplete emptying, and Nocturia), was developed at our institution to reduce the burden of the AUA-SS questionnaire. In addition to these 4 domains, the UWIN also contains a QOL index question. The UWIN is drastically shorter than the AUA-SS because it consists of just 4 questions, with scores ranging from 0-3, adding up to a composite score of 0-12.

To create the UWIN, we used results of a study of 8731 participants at Prostate Cancer Awareness Week in 2003 and 2004 who completed the AUA-SS. We excluded questions on frequency, hesitancy, and intermittence, and reduced the range from 6 answers to 4. The receiver operating characteristic curve based on survey answers placed patients in 3 categories: mild (0-3), moderate (4-8), and severe (9-12).[5] Next, we evaluated the UWIN with 278 attendees of Prostate Cancer Week in 2006 and 2007. Bland-Altman plots demonstrated agreement between the AUA-SS and UWIN, and the Spearman coefficient was 0.931 (P <.0001), indicating a strong relationship between them.[6]

MATERIALS AND METHODS

This prospective, cohort study of the efficacy of the UWIN study in assessing LUTS compared with that of the AUA-SS was approved by the Colorado Multiple Review Board-approved (11-1625).

Study Setting and Population

Patients were enrolled in 2011-2012 at the urology clinic at the University of Colorado Hospital. All patients attending the clinic were recruited for this study to establish generalizability among patients with LUTS and increase sample size for better power for comparison. Enrollment criteria stated patients were required to complete both surveys on the same day.

Data Collection

We distributed randomly the UWIN and AUA-SS among 700 patients, and 593 completed both surveys. The response rate was 84.7%. Among these patients, 26.1% were preoperative, 11.1% were postoperative, and 62.7% had diseases other than BPH. Demographic data were also collected from the patients.

Outcome Measures

The UWIN and AUA-SS total scores, QOL, and categories were study outcomes to detect the agreement between them.

Table 1. Baseline characteristics of 593 patients

Baseline Characteristics	Mean ± SD or Median (IQR)
Age (y)	66 ± 10
AUA-SS score	9 (4, 17)
UWIN score	4 (2, 7)

AUA-SS, American Urology Association symptom score; IQR, interquartile range; SD, standard deviation; UWIN, Urgency, Weak stream, Incomplete emptying, and Nocturia survey.

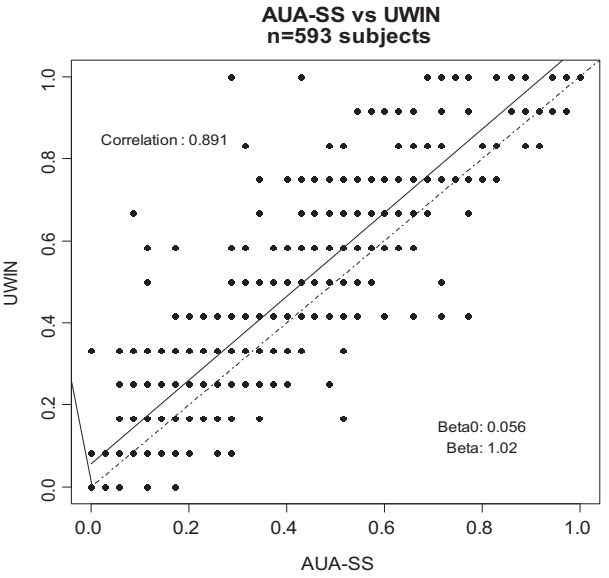

Figure 1. Spearman correlation between the American Urology Association Symptom Score (AUA-SS) and the Urgency, Weak stream, Incomplete emptying, and Nocturia survey (UWIN).

Statistical Analysis

AUA-SS and UWIN survey results, age, and race were considered descriptive statistics. Because the AUA-SS is scored from 0-5 and the UWIN from 0-3, we scaled the scored to a range of 0-1 to equate the measures. This was done by dividing the score by the maximum total UWIN score of 12 or the maximum total AUA-SS score of 35. The relationship among the surveys was determined using the Spearman correlation coefficient and Bland-Altman plots for agreement.[7] All analysis was conducted using SAS 9.2 software (SAS Institute).

RESULTS

This study included 593 patients who were white (32.34%), African American (3.56%), Asian (1.34%), other (3.41%), unknown (58.9%), Pacific Islander (0.15%), and refused to answer (0.30%). Table 1 lists further information about patients' demographics.

The Spearman correlation for the total AUA-SS and UWIN was 0.891, which was statistically significant (P <.001; Fig. 1). At the level of individual questions, the correlation coefficients between the 2 tests were 0.82 for urgency, 0.89 for weak urinary stream, 0.85 for incomplete emptying, and 0.81 for nocturia (P <.001). The Spearman correlation coefficient for the QOL

(ie, bothersome question) was 0.87 and statistically significant ($P < .001$). We also assessed the categories. The Spearman correlation coefficient between AUA-SS categories and UWIN categories was 0.81, which was significant ($P < .001$). The Spearman correlation coefficient between the AUA-SS and the UWIN at the level of total score, category, and QOL showed a strong positive relationship (Table 2).

The Bland-Altman plots between total AUA-SS and UWIN (Fig. 2) and between categories of AUA-SS and UWIN shows good agreement. This indicates that even after deleting items and changing answer options, the UWIN demonstrates a similar degree of utility as the AUA-SS. This supports the notion that the shorter UWIN survey can be used in place of the longer AUA-SS. A high level of agreement between QOL in AUA-SS and UWIN is demonstrated by the Bland-Altman graph indicating that decreasing the non-QOL answer options from 7 to 3 did not alter the value of the survey.

COMMENT

Previous research shows that there is a significant problem regarding the degree to which patients misunderstand items on the AUA-SS, a globally used survey to evaluate LUTS. To resolve this problem, a shorter, modified version of the AUA-SS was developed termed the UWIN.[6-9]

A 2009 prospective study of 407 patients from a county hospital and a university hospital compared answers from interviewer-administered and self-administered AUA-SS and found that only 72% understood all 7 questions. Of the measured demographic variables, only education level significantly affected this understanding. Compared with patients with more than 12 years of education, county hospital patients with less than 9 years of education were 57.06 times more likely to misunderstand AUA-SS questions (95% confidence interval, 14.32-329.34), whereas university hospital patients with less than 9 years of education were 38.27 times more likely to misunderstand AUA-SS questions (95% confidence interval, 1.69-867.83). Misrepresented complaints as a result of misunderstood questions occurred in 21% of university hospital patients and in 31% of county hospital patients. Patients with low education, regardless of location, were more prone to misunderstand the AUA-SS survey, misrepresent their symptoms, and ultimately, receive inappropriate treatment.[8]

Some patients may experience difficulty completing the AUA-SS due to the wording or length of the questions, or both. A 2008 study of 989 patients who self-administered the survey and completed it with the aid of a physician found that higher education levels correlated with less misrepresentation of symptoms. This problem was obvious for the frequency item and the urgency item. They found that 28% of patients with more than 12 years of education were misrepresented, whereas

Table 2. Spearman correlation coefficients between American Urology Association Symptom Score and Urgency, Weak stream, Incomplete emptying, and Nocturia survey at the level of individual questions for 593 patients

		AUA									UWIN						Category	
$P < .001$ for all		Score	IE	AUA2	AUA3	U	W	AUA6	N	QOL	Score	U	W	IE	N	QOL	AUA	UWIN
AUA	Score	1.00																
	IE	0.79	1.00															
	AUA2	0.80	0.57	1.00														
	AUA 3	0.77	0.63	0.53	1.00													
	U	0.74	0.50	0.58	0.45	1.00												
	W	0.80	0.60	0.47	0.63	0.50	1.00											
	AUA6	0.69	0.56	0.43	0.60	0.42	0.63	1.00										
	N	0.64	0.36	0.56	0.35	0.44	0.36	0.30	1.00									
	QOL	0.79	0.60	0.64	0.53	0.54	0.59	0.49	0.53	1.00								
UWIN	Score	0.90	0.73	0.67	0.64	0.70	0.77	0.61	0.59	0.74	1.00							
	U	0.70	0.51	0.59	0.44	0.82	0.48	0.42	0.44	0.56	0.78	1.00						
	W	0.74	0.56	0.43	0.59	0.46	0.89	0.62	0.33	0.57	0.80	0.49	1.00					
	IE	0.75	0.85	0.55	0.63	0.48	0.60	0.57	0.32	0.61	0.80	0.52	0.62	1.00				
	N	0.56	0.33	0.43	0.31	0.36	0.32	0.28	0.81	0.46	0.63	0.40	0.31	0.33	1.00			
	QOL	0.70	0.54	0.57	0.47	0.52	0.56	0.48	0.46	0.87	0.70	0.56	0.55	0.57	0.44	1.00		
Cat	AUA	0.92	0.75	0.75	0.71	0.68	0.74	0.65	0.58	0.71	0.83	0.64	0.69	0.72	0.50	0.64	1.00	
	UWIN	0.84	0.71	0.63	0.61	0.65	0.72	0.59	0.53	0.69	0.93	0.73	0.75	0.79	0.55	0.66	0.81	1.0

AUA2, AUA-SS symptom 2; AUA3, AUA-SS symptom 3; AUA6, AUA-SS symptom 6; IE, incomplete emptying, N, nocturia; QOL, quality of life; U, urgency; W, weak urinary stream; other abbreviations as in Table 1.

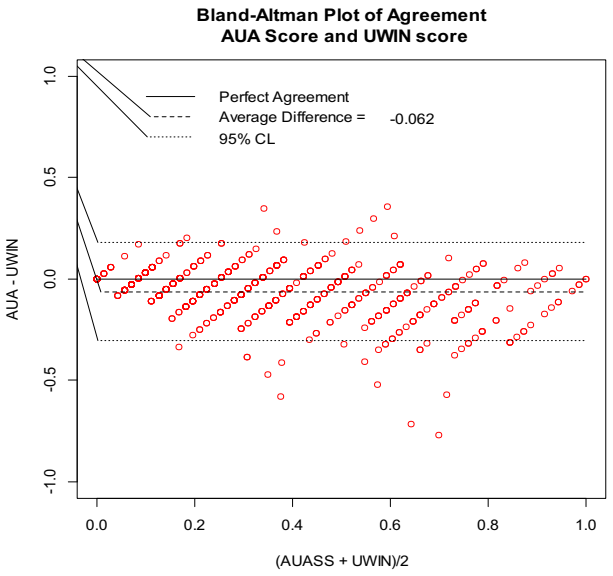

Figure 2. Bland-Altman graph shows agreement between the American Urology Association Symptom Score (AUA-SS) and the Urgency, Weak stream, Incomplete emptying, and Nocturia survey (UWIN). (Color version available online.)

58% with fewer than 9 years of education were misreported.[9] The authors concluded that a shorter, simpler survey could help to improve these limitations.[10]

In a randomized, controlled trial in 2009, 232 patients were assigned to the traditional self-administered AUA-SS or a new multimedia version of the survey as an experimental arm. Patients in both arms were administered the survey a second time for comparison. An observed 43% decline in error was seen in the experimental arm. No association was found between improvements and patient education level. The AUA-SS multimedia version represents an excellent opportunity to improve the AUA-SS as a screening tool because it was shown to increase understanding and decrease scoring errors across all education levels, possibly allowing physicians to treat patients more effectively.[11]

In this study, the Spearman correlation coefficient showed a strong positive correlation between the AUA-SS and UWIN. Good agreement on Bland-Altman plots indicate that, although shorter and easier to complete, the UWIN does not detract from the validity and reliability of the AUA-SS. Our results show that the UWIN survey may replace the AUA-SS because of the strong positive relationship and agreement between them.

One of the major limitations of this study was the patients' ability and willingness to answer survey questions. Some patients asked others to complete the survey on their behalf, and others faced problems with answering, leading them to miss questions, select multiple choices, or write descriptive notes in lieu of choosing an appropriate answer. We resolved this problem by instructing all patients to complete the survey independently to the best of his ability.

Future studies of the UWIN questionnaire will include urodynamic workup, and consideration will be given for the evaluation of AUA-SS and UWIN scoring as they relate to various BPH treatments. A test-retest will be done for the UWIN survey. We will also ask respondents their opinions on which survey was easier to understand.

CONCLUSIONS

Misunderstanding and misinterpretation are major limitations of using the AUA-SS. We developed the UWIN survey to solve these problems. Here, we have presented data that patient responses to the UWIN and AUA-SS are similar, and thus, this shorter and easier questionnaire may potentially replace the AUA-SS.

Acknowledgments. The authors thank John D. Jones, Jr., M.S.I.S., Associate Professor and Librarian at Health Sciences Library in University of Colorado, for his help and assistance during the literature review.

References

1. Black L, Naslund MJ, Gilbert TD Jr, et al. An examination of treatment patterns and costs of care among patients with benign prostatic hyperplasia. *Am J Manag Care*. 2006;12:S99-S110.
2. McNaughton-Collins M, Barry MJ. Managing patients with lower urinary tract symptoms suggestive of benign prostatic hyperplasia. *Am J Med*. 2005;118:1331-1339.
3. Roehrborn CG, McConnell JD, Barry MJ et al. AUA Guideline on the Management of Benign Prostatic Hyperplasia. Chapter 1-5 through 1-12. 2003 (updated 2010). Available at: http://www.auanet.org/common/pdf/education/Arc-BPH-Chapter1.pdf. Accessed June 28, 2013.
4. Tanguay S, Awde M, Brock G, et al. Diagnosis and management of benign prostatic hyperplasia in primary care. *Can Urol Assoc J*. 2009; 3:S92-S100.
5. Barqawi AB, Sullivan KF, Crawford ED, et al. Methods of developing UWIN, the modified American Urological Association symptom score. *J Urol*. 2011;186:940-944.
6. Crawford ED, O'Donnell CI, Barqawi AB, et al. Validation of the modified American Urological Association symptom score. *J Urol*. 2011;186:945-948.
7. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327: 307-310.
8. Timothy V, Evan D, Abbasi A, et al. Assessment of the performance of the American Urological Association Symptom Score in 2 distinct patient populations. *J Urol*. 2009;181:230-237.
9. Johnson TV, Abbasi A, Ehrlich SS, et al. Patient misunderstanding of the individual questions of the American Urological Association symptom score. *J Urol*. 2008;179:2291-2295.
10. Coste J, Guillemin F, Pouchot J, et al. Methodological approaches to shortening composite measurement scales. *J Clin Epidemiol*. 1997; 50:247-252.
11. Bryant MD, Schoenberg ED, Johnson TV, et al. Multimedia version of a standard medical questionnaire improves patient understanding across all literacy levels. *J Urol*. 2009;182:1120-1125.

APPENDIX

UWIN Symptom Index©

Name: _____Date: _____

DOB: _____MRN: _____

Question	Not At All	Less Than Half the Time	About Half the Time	More Than Half the Time
Urgency: Over the past month or so, how often have you found it difficult to postpone urination?	0	1	2	3
Weak Stream: Over the past month or so, how often have you had a weak urinary stream?	0	1	2	3
Incomplete Emptying: Over the past month or so, how often have you had a sensation of not emptying your bladder completely after you finished urinating?	0	1	2	3
Nocturia: Over the last month, how many times did you most typically get up to urinate from the time you went to bed at night until the time you got up in the morning?	None 0	1-2 Times 1	3 Times 2	4 or More Times 3
Bothersome: If you were to spend the rest of your life with your urinary condition just the way it is now, how would you feel about that?		Happy 0	Mixed 1	Unhappy 2