The Comprehensive Assessment of Self-Reported Urinary Symptoms (CASUS): A New Tool for Research on Subtypes of Patients with Lower Urinary Tract Symptoms

Kevin P. Weinfurt, PhD[1], **James W. Griffith, PhD**[2], **Kathryn E. Flynn, PhD**[3], **David Cella, PhD**[2], **Tamara Bavendam, MD**[4], **Jonathan B. Wiseman, MS**[5], **Victor P. Andreev, PhD, DSc**[5], **H. Henry Lai, MD**[6], **Alice B. Liu, MPH**[7], **Ziya Kirkali, MD**[4], **Anne P Cameron, MD**[8], **Catherine S. Bradley, MD, MSCE**[9], and **LURN Study Group**

[1]Duke University Medical Center, Durham, NC

[2]Northwestern University, Feinberg School of Medicine, Department of Medical Social Sciences, Chicago, IL

[3]Medical College of Wisconsin, Milwaukee, WI

[4]National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD

[5]Arbor Research Collaborative for Health, Ann Arbor, MI

[6]Washington University School of Medicine, St. Louis, MO

[7]University of Washington, Seattle, WA

Corresponding Author: Kevin P. Weinfurt, PhD, Professor and Vice Chair of Research, Department of Population Health Sciences, Duke University School of Medicine, 215 Morris, Street Box 104023, Durham, NC 27705, Tel: (919) 668-8019, Kevin.Weinfurt@duke.edu.

[8]University of Michigan, Ann Arbor, MI

[9]University of Iowa, Carver College of Medicine, Department of Obstetrics and Gynecology, Iowa City, IA

Abstract

Purpose: To improve the potential for finding clinically important subtypes of patients with lower urinary tract symptoms (LUTS), we describe the development of the Comprehensive Assessment of Self-reported Urinary Symptoms (CASUS) —and use it to present data on the experiences of LUTS in treatment-seeking women and men from a prospective observational cohort.

Materials and Methods: An initial list of LUTS as confirmed in 22 qualitative interviews with providers and 88 qualitative interviews with care-seeking and non-care-seeking women and men with LUTS. Items from extant measures were adopted and revised and new items were developed, and all were evaluated for understanding in 64 cognitive interviews. Items were administered to a prospective cohort of female and male LUTS patients seeking care and analyses were conducted to describe item response distributions and correlations among item responses separately for women and men.

Results: A total of 444 males and 372 females provided responses to CASUS. There were several sets of items that had different relationships for women compared to men. In particular, the associations between sensation-related items and incontinence-related items were generally positive among females, but were often negative among males.

Conclusions: Using an intensive development process, the CASUS addresses a wide range of LUTS. It should help to identify clinically important subtypes of patients. Further, the collection of items can provide the foundation for shorter measures for use in the clinic and as trial endpoints.

Keywords

Introduction

The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN) is working to identify clinically relevant subtypes of patients with lower urinary tract symptom (LUTS).[1] To conduct this research, LURN used diverse approaches to characterizing patients with LUTS by assessing patients' reported experience, neurological and sensory factors, lower urinary tract organ structure and function, and potential biomarkers.

Although self-reported measures of LUTS have been developed,[2–4] they fail to assess certain concepts and are not comprehensive for both men and women. Thus, LURN sought to create a comprehensive questionnaire appropriate for both women and men, containing a wide range of symptoms, and that might provide granular and precise information about symptom experience. Additionally, we wished to explore the boundaries of the types of LUTS experiences that could be self-reported, especially with respect to unique or abnormal

bladder sensations. We also sought to create a pool of items from which researchers could draw to create additional clinical tools and outcomes measures. We describe the multi-phase development of the measure—the Comprehensive Assessment of Self-reported Urinary Symptoms (CASUS) —and use it to present data on the experiences of LUTS in treatment-seeking women and men from a prospective observational cohort.

Methods

Table 1 provides an overview of the CASUS development process, which followed the same general approach used in the National Institutes of Health' Patient-Reported Outcomes Measurement Information System.[5,6,7] A list of LUTS based on literature review and expert opinion from LURN investigators was developed. The initial working list included storage, voiding, and post-micturition symptoms, as well as related concepts including confidence in warning signs of need to urinate, self-rating of overall bladder control, urgency with fear of leaking, abnormal bladder sensations, and symptom bother. This list both guided and was refined by subsequent development steps.

Qualitative Interviews with Providers

To understand providers' perspectives on the most important aspects of LUTS, individual semi-structured qualitative interviews were conducted.

Sample.—We recruited 15 physicians for interviews—10 with a urology or urogynecology specialty (five treating women and five treating men) and five in primary care. All physicians had to be board certified, have more than five years of experience, and evaluate more than five patients with LUTS per week. Six nurses were also interviewed—four who worked primarily with men and two who worked primarily with women. All providers were recruited from professional networks of LURN investigators to reflect geographic diversity.

Procedure.—Semi-structured interviews were conducted by a trained research assistant either in-person or by telephone. Participating clinicians were asked to list and identify the most important and prevalent LUTS and concerns using open-ended queries, as well as to suggest patient-friendly language for each of the symptoms discussed.

Outcome.—The provider interviews confirmed that the working symptom list was acceptable.

Qualitative Interviews with Patients and Community Members

We conducted qualitative interviews with patients and community-dwelling women and men who report LUTS to evaluate our working symptom framework and to hear how men and women with LUTS describe their experiences in their own words. We also sought to understand the experiences of people who have abnormal bladder sensations, as this was identified as an area that had not been assessed well in extant instruments. Details of the methods are described in Griffith et al.[6]

Sample.—Participants had to report one or more LUTS including storage, voiding, and post-micturition symptoms. We sought to recruit at least 30 patients who had sought care for

LUTS and at least 30 community members with LUTS who had not sought care, with equal proportions of males and females. Patients were recruited from four LURN sites and community members were recruited via advertisements (e.g., Craigslist™). Additionally, we sought to recruit at least 16 patients (half males, half females) who were likely to have abnormal bladder sensations, or lack of sensation—namely, those with a recent lower spinal cord injury, recent lower back surgery, women with a recent difficult vaginal delivery, women with a recent radical hysterectomy, underactive bladder, people with uncontrolled diabetes, and older individuals (age 65+). A minimum of 25% of all participants were required to be non-White.

Procedure.—The protocol was approved by the institutional review boards at the participating recruitment sites and the LURN Data Coordinating Center. All participants provided written informed consent. Participants had an in-person interview with a trained staff member. This qualitative interview addressed questions about specific urinary symptoms the person experienced, including onset, duration, and degree of bother. The participant was also asked to describe what made the symptoms better or worse, and what they did to cope with their symptoms. Finally, participants were asked whether they had sought care for their LUTS, and reasons why or why not (c.f. Griffith et al.[6]). Following the interview, each participant completed the LUTS Tool[3] to assess symptoms and bother using a standardized approach.

Analysis.—Transcripts were created from audio recordings of each interview, and two trained raters assigned qualitative codes to identify key themes.

Outcome.—Eighty-eight people with LUTS participated, including 34 recruited from clinics, 33 from the community, and 21 at risk for abnormal bladder sensations. The sample was 57% female and 77% White, with mean age of 52 (range 19 to 77). Findings from the qualitative interviews confirmed a range of symptoms in multiple areas, including storage, voiding, and post-voiding. In addition, some individual symptoms were experienced in a variety of ways. For example, some participants reported a lack of sensation with voiding and leakage whereas for other participants, leakage was associated with urge sensations or urgency.

Collecting and Writing Items

Using the symptom framework that was verified through the preceding qualitative studies, we developed a pool of candidate items drawing on existing items from the LUTS Tool, the AUA Symptom Index (AUA-SI), Sung et al[8]., as well as creating novel items. Items were binned into symptom dimensions (e.g., nocturia) and pairs of LURN investigators reviewed existing items to flag those that were potentially irrelevant or redundant. Modifications were made where necessary, and new items were drafted using item-writing conventions based on experience in PROMIS (Patient-Reported Outcome Measurement Information System) and other measure development efforts. These conventions included using second person, past tense, and "in the past 7 days" as the default recall period, consistent with most PROMIS scales, but under active evaluation in a separate LURN protocol. The resulting lists of items for each dimension were reviewed by a third LURN investigator, and then presented to the

LURN Self-Reported Measures workgroup for review. This workgroup reviewed and approved 68 items for inclusion in the next phase of testing.

Cognitive Interviews and Item Revision

To ensure that items were understood as intended and that respondents were able to choose a response that matched their experience, we conducted one-on-one cognitive interviews[9]. Items were then revised and retested in another round of cognitive interviews, as needed.

Sample.—The sample consisted of people with and without LUTS. Participants with LUTS were recruited from LURN-affiliated clinics and through community advertising as described earlier. Participants without LUTS were recruited through community advertising. Within both LUTS and non-LUTS groups, we targeted equal numbers of females and males.

Procedure.—The protocol was approved by the institutional review boards at the participating recruitment sites and the Data Coordinating Center. After providing informed consent, each participant was presented with a maximum of 35 of the 68 CASUS items in a face-to-face interview. The participant was asked to complete each item and then answer questions that addressed comprehension, appropriateness of response options and recall period, and how the participant arrived at his or her response. Participants were asked for suggestions for improving items. Interviewers recorded observations about each participant's reaction to each item, which were later reviewed by the Self-Reported Measurement workgroup to inform item revisions. Following review of all items, each participant was administered the Wide Range Achievement Test (WRAT) Reading subtest to assess literacy level[8].

Each item was reviewed initially by five men and five women with LUTS, and five men and five women without LUTS. Of the 10 men and 10 women viewing each item, two to three men and women had to have lower literacy, defined as a reading level less than ninth grade using the WRAT-4 Reading subtest or less than twelfth grade education or equivalent. We also required that each item be reviewed by at least two participants with LUTS endorsing the target symptom, and that every item be reviewed by at least one white and one non-white person with LUTS and without LUTS.

Outcome.—Across two rounds of cognitive interviews, 53 participants with LUTS and 11 without LUTS participated. Table 2 describes their demographic characteristics. Substantial revisions were needed for several items, which necessitated a second round of cognitive interviews.

Translatability Review

Translation into additional languages is beyond the current scope of LURN, but to facilitate future translations, items were submitted to a translatability review by experts from Northwestern University. Minor wording changes were required for some items to facilitate easier translation into a wide variety of languages. The final set of CASUS items is shown in Appendix A.

Statistical Evaluation of Items

To assess the distribution of CASUS symptom responses in a broad population of LUTS patients, we included the CASUS as part of the LURN Observational Cohort, a prospective study of patients presenting with LUTS at the six LURN sites[11]. The objective of this quantitative study was to describe the item response distributions and inter-item associations in men and women.

Sample.—The methods for recruitment and data collection are described in detail in Cameron et al[11]. Briefly, women and men were recruited from urology and urogynecology clinics at participating LURN sites.

Procedure.—Participants completed the CASUS items electronically during their baseline, and 12-month study visits. CASUS items were added to the LURN Observational Cohort study after it began enrolling, so relatively few had CASUS items at baseline (N = 287). Accordingly, we focused on CASUS data collected at the 12-month visit. The protocol was approved by the institutional review boards at the participating recruitment sites and the Data Coordinating Center.

Statistical Analyses.—Demographics and medical history of participants completing CASUS at their 12-month visit are presented by sex, with means and standard deviations (or medians and interquartile ranges, where appropriate) for continuous variables, and frequencies and percentages for categorical variables. Differences in sex were assessed using chi-square and Wilcoxon two-sample tests. CASUS item response frequencies and percentages were computed. Polychoric correlations were calculated for all non-sex-specific pairs of CASUS items and are presented in correlation heat map matrices separately by sex. Differences in the correlations between sexes are presented in a separate heat map. To determine whether differences were statistically significant, we used a bootstrap method[12]. This analysis was performed on all CASUS items as well as each section of CASUS. All statistical tests were performed using SAS Version 9.4 (Cary, North Carolina).

Results

Table 3 summarizes LURN Observational Cohort participant characteristics at the 12-month study visit. Supplementary Table 1 shows the distribution of CASUS item responses by sex. Most of the items demonstrated a range of responses across the response categories, though most had non-uniform distributions. Certain items had relatively low endorsements in this sample, including items about leaking at night (B4 and B5) and several of the items addressing sensations experienced when the respondent felt the need to urinate (C1a-C3m). Of the latter sensations, seven were endorsed by   10% of respondents: sensations in scrotum/testicles (C1e), sensations in the lower back (C1g and C2e), burning (C3d), pain (C3g), aching (C3h), fullness (C3j), and bloating (C3l).

Graphical depictions of the correlation heat map matrices of the responses to CASUS questions are presented for females (Figure 1), males (Figure 2), and the difference between them (Figure 3). As Figure 3 shows, there were several sets of items that have different relationships for women compared to men, and the overall correlation heat map matrices for

all CASUS items were shown to be statistically significantly different between sexes (estimate p =.0004). In particular, the associations between sensation-related items and incontinence-related items were generally positive among females, but were often negative among males. (The estimated probability for this region alone was.0002.)

Discussion

The ability to identify clinically meaningful subgroups of patients depends on the granularity, accuracy, and reliability of the data. In an effort to improve the quality of self-reported symptom data available for such research, LURN developed the CASUS. The strengths of the development process included intensive engagement with health care providers who treat LUTS, women and men with LUTS who did and did not seek care for their LUTS, plus community individuals without LUTS. We also included aspects of symptoms not covered by any single measure currently in use (e.g., different aspects of urinary incontinence, dribbling and post-void symptoms, and abnormal bladder sensations before and between urination[13]), careful attention to the wording of items and their evaluation through cognitive interviews, which included participants with lower levels of literacy, a translatability review to facilitate later translations into other languages, and helpful reference data from over 816 patients seeking care for LUTS.

To increase the granularity of information available for later research, efforts were made to create items that address different nuances of symptoms, including different sensory qualities. The inclusion of these sensory qualities has already helped us to observe how different sensations can co-occur with incontinence in women, but less so for men. Future work is needed to verify this pattern of sex differences and to identify potential causal mechanisms. Other plans include using the highly-granular symptom data to predict treatment outcomes and to understand etiological mechanisms. Finally, although the CASUS was created to be used in its entirety to discover clinical subtypes of LUTS, LURN is also developing brief versions of CASUS for use in clinical care and research.

There are several limitations in this work. First, the CASUS was developed in English only, but our translatability review sets a strong foundation for future translations. Second, our recruitment of patients with LUTS was restricted to specialty clinics within academic medical centers. Thus, the generalizability of findings to patients seen in other settings is unknown. We tried to reduce this concern by including community members who had not sought care in our qualitative work, as well including participants without LUTS in our cognitive interviews. Finally, we intended to elicit more experiences of abnormal/reduced bladder sensations by conducting qualitative work with patients thought to be at risk for such experiences. Our qualitative interview participants did not report abnormal experiences as much as we had expected and so future qualitative work is needed to better understand abnormal urinary and bladder sensations.

Conclusions

Using an intensive development process, the CASUS addresses a wide range of LUTS. Combined with other types of clinical and lab data, it should be useful in future LUTS

research, as well as help to identify clinically important subtypes of patients. Further, the exhaustive collection of items can provide the foundation for shorter measures for use in the clinic and as trial endpoints.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Yang CC, Weinfurt KP, Merion RM, et al. Symptoms of Lower Urinary Tract Dysfunction Research Network. J Urol. 196:146–152, 2016 [PubMed: 26784646]

2. 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.148:1549–1557, 1992 [PubMed: 1279218]

3. Coyne KS, Barsdorf AI, Thompson C et al. Moving towards a comprehensive assessment of lower urinary tract symptoms (LUTS). Neurourol Urodyn. 31:448–454, 2012 [PubMed: 22396308]

4. Abrams P, Avery K, Gardener N, Donovan J; ICIQ Advisory Board. The International Consultation on Incontinence Modular Questionnaire: www.iciq.net. J Urol. 175: 1063–1066, 2006 [PubMed: 16469618]

5. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes. 4:79, 2006 [PubMed: 17034633]

6. Griffith JW, Messersmith EE, Gillespie BW et al. Reasons for Seeking Clinical Care for Lower Urinary Tract Symptoms: A Mixed Methods Study. J Urol. 199: 528, 2018 [PubMed: 28734864]

7. DeWalt DA, Rothrock N, Yount S, et al. Evaluation of item candidates: the PROMIS qualitative item review. Medical Care. 45(Suppl 1):S12–S21, 2007. [PubMed: 17443114]

8. Sung VW, Griffith JW, Rogers RG, Raker CA, Clark MA. Item bank development, calibration and validation for patient-reported outcomes in female urinary incontinence. Qual Life Res. 25: 1645–1654, 2016 [PubMed: 26732514]

9. Willis GB. Cognitive Interviewing: A Tool for Improving Questionnaire Design: Sage Publications Inc., Thousand Oaks, CA, USA 2004

10. Wilkinson GS, Robertson GJ. WRAT4: Wide Range Achievement Test: Psychological Assessment Resources, 2006

11. Cameron AP, Lewicky-Gaupp C, Smith AR, et al. Baseline Lower Urinary Tract Symptoms in Patients Enrolled in the Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN): a Prospective, Observational Cohort Study. J Urol. 199:1023–1031, 2018. [PubMed: 29111381]

12. Roff DA, Mousseau TA, Howard DJ.Variation in Genetic Architecture of Calling Song among Populations of Allonemobius socius, A. fasciatus, and a Hybrid Population: Drift or Selection? Evolution, 53:216–224, 1999 [PubMed: 28565178]

13. Helfand BT, Smith AR, Lai HH, et al. Prevalence and Characteristics of Urinary Incontinence in a Treatment Seeking Male Prospective Cohort: Results from the LURN Study. J Urol. 200:397–404, 2018 [PubMed: 29477718]

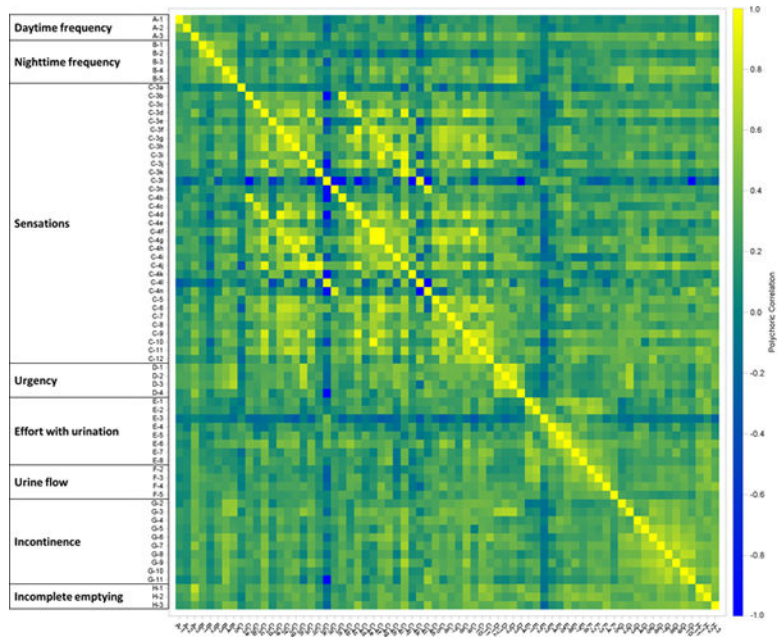**Figure 1.**
Pairwise polychoric correlations of CASUS items (Females only, sex-specific questions removed).
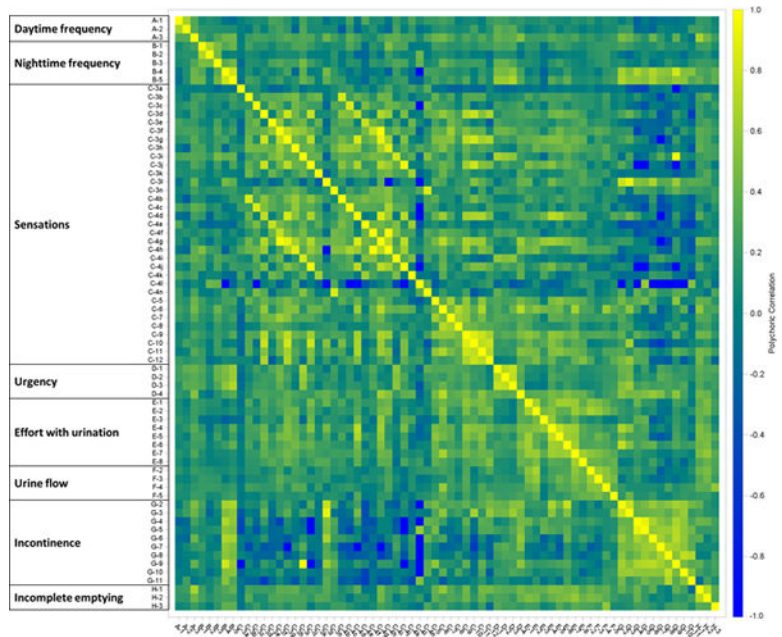
Figure 2.
Pairwise polychoric correlations of CASUS items (Males only, sex-specific questions removed).
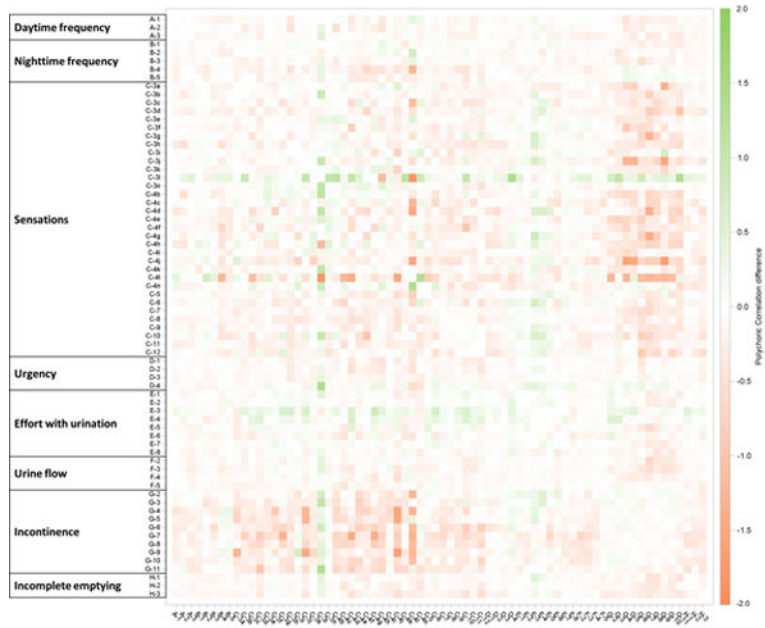
Figure 3.
Pairwise polychoric correlations of CASUS items (Differences between males and females, sex-specific questions removed) *Note: Green color indicates males have a higher correlation than females; orange indicates females have a higher correlation than males.*

Table 1.

Overview of CASUS Development Process

Step	Sample Size
1. Qualitative interviews with clinical experts	22
2. Qualitative interviews with people with LUTS	88
3. Identify existing items and write new/improved items	NA
4. Cognitive interviews to check for understanding and appropriateness	64
5. Translatability review	---
6. Determine best recall period	517
7. Statistical analysis to select final items	816

Table 2.

Participant Characteristics for the Cognitive Interviews

	LUTS (n=53)	Non-LUTS (n=11)
Age (yrs.)	41.2 (14.7)	55.5 (18.3)
Sex (Male)	23 (43%)	4 (36%)
Race		
American Indian/Alaskan Native	0 (0%)	0 (0%)
Asian	2 (4%)	1 (9%)
African-American	17 (32%)	2 (18%)
Native Hawaiian/Pacific Islander	1 (2%)	0 (0%)
White	31 (58%)	8 (73%)
Multi-racial/Other	2 (4%)	0 (0%)
Lower Literacy[a]	10 (19%)	5 (45%)

[a]Defined as a reading level less than ninth grade using the WRAT-4 Reading subtest or less than twelfth grade education or equivalent (e.g., GED).

Table 3.

Demographics and Medical History of LURN Participants by Sex

	Male (n=444)	Female (n=372)	Total (n=816)	p-value[*]
Age	62.0 (12.6)	57.4 (13.8)	59.9 (13.4)	<.001
Race				0.437
American Indian/Alaskan Native	2 (0%)	3 (1%)	5 (1%)	
Asian	18 (4%)	10 (3%)	28 (4%)	
African-American	47 (11%)	50 (14%)	97 (12%)	
Native Hawaiian/Pacific Islander	0 (0%)	1 (0%)	1 (0%)	
White	355 (83%)	301 (82%)	656 (83%)	
Multi-racial/Other	5 (1%)	2 (1%)	7 (1%)	
BMI	29.6 (5.5)	30.6 (7.7)	30.1 (6.6)	0.565
Current or Former Smoker	206 (47%)	125 (66%)	331 (41%)	<.001
Diabetes	77 (17%)	53 (14%)	130 (16%)	0.250
Number of Vaginal Births	-	1.8 (1.5)	-	
Post-menopausal	-	255 (69%)	-	
Hormone Use	-	47 (18%)	-	
Anticholinergic medication use	15 (3%)	8 (2%)	23 (3%)	0.396
Constipation medication use	38 (9%)	25 (7%)	63 (8%)	0.358
Previous Surgery (multiple possible)				
Urge incontinence	1 (0%)	3 (1%)	4 (0%)	0.336
SUI/Prolapse	-	49 (13%)	-	
Hysterectomy	-	116 (31%)	-	
Prostate	22 (5%)	-	-	
Urethral Dilation	0 (0%)	3 (1%)	3 (0%)	0.094
Other	0 (0%)	2 (1%)	2 (0%)	0.208
Functional Comorbidity Index	2.3 (1.8)	2.4 (2.2)	2.3 (2.0)	0.867
Prolapse Stage (n=317)				
Stage 0	-	94 (30%)	-	
Stage 1	-	106 (33%)	-	
Stage 2	-	93 (29%)	-	
Stage 3	-	23 (7%)	-	
Stage 4	-	1 (0%)	-	
Prostate Findings (n=382)				
Nodule/Anomaly	8 (2%)	-	-	
Normal/enlarged prostate	374 (98%)	-	-	
PVR (ml) (n=347 male, 318 female)	70.6 (115.6)	44.3 (59.7)	58.0 (94.0)	0.143
AUA-SI (n=397 male, 339 female)	13.7 (7.1)	12.1 (6.1)	13.0 (6.7)	0.003
AUA QOL (n=390 male, 335 female)	3.7 (1.4)	4.4 (1.3)	4.0 (1.4)	<.001

[*] P-value for male vs. female from chi-square test or Wilcoxon 2-sample test