Shauna Leighton ORCID iD: 0000-0001-5700-1424

Claire Yang ORCID iD: 0000-0002-4750-6258

H. Henry Lai ORCID iD: 0000-0003-2691-994X

Megan Bradley ORCID iD: 0000-0001-9935-5132

Are Three-Day Voiding Diaries Feasible and Reliable? Results from the Symptoms of Lower Urinary Tract Dysfunction Network (LURN) Cohort

Anne P. Cameron[1], Jonathan B. Wiseman[2], Abigail R. Smith[2], Robert M. Merion[2], Brenda W. Gillespie[1], Catherine S. Bradley[3], Cindy L. Amundsen[4], Claire C. Yang[5], H. Henry Lai[6], John O.L. DeLancey[1], Margaret E. Helmuth[2], Megan S. Bradley[4], Nnena Agochukwu[1], Victor P. Andreev[2], Ziya Kirkali[7], J. Quentin Clemens[1], and the LURN Study Group

[1]University of Michigan, Ann Arbor, MI

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

[3]University of Iowa Carver College of Medicine, Iowa City, IA

[4]Duke University Medical Center, Durham, NC

[5]University of Washington, Seattle, WA

This is the author manuscript accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/nau.24113.

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

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

Corresponding Author:

Anne P. Cameron

1500 East Medical Center Drive

Taubman Center 3875, Ann Arbor, MI, 48109-5330, USA

Phone: 734-936-7030; Fax: 734-232-1610 Email: annepell@med.umich.edu

Short title: Reliability of three-day voiding diary

Word count: 2496

Declarations of Interest: none

Funding/Support

This is publication number 19 of the Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN).

This study is supported by the National Institute of Diabetes & Digestive & Kidney Diseases through cooperative agreements (grants DK097780, DK097772, DK097779, DK099932, DK100011, DK100017, DK097776, DK099879).

Research reported in this publication was supported at Northwestern University, in part, by the National Institutes of Health's National Center for Advancing Translational Sciences, Grant Number UL1TR001422. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Dr. Siddiqui is supported by grant K23-DK110417 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

Acknowledgments

The following individuals were instrumental in the planning and conduct of this study at each of the participating institutions:

Duke University, Durham, North Carolina (DK097780): PI: Cindy Amundsen, MD, Kevin Weinfurt, PhD; Co-Is: Kathryn Flynn, PhD, Matthew O. Fraser, PhD, Todd Harshbarger, PhD, Eric Jelovsek, MD, Aaron Lentz, MD, Drew Peterson, MD,

Nazema Siddiqui, MD, Alison Weidner, MD; Study Coordinators: Carrie Dombeck, MA, Robin Gilliam, MSW, Akira Hayes, Shantae McLean, MPH

University of Iowa, Iowa City, IA (DK097772): PI: Karl Kreder, MD, MBA, Catherine S Bradley, MD, MSCE, Co-Is: Bradley A. Erickson, MD, MS, Susan K. Lutgendorf, PhD, Vince Magnotta, PhD, Michael A. O'Donnell, MD, Vivian Sung, MD; Study Coordinator: Ahmad Alzubaidi

Northwestern University, Chicago, IL (DK097779): PIs: David Cella, Brian Helfand, MD, PhD; Co-Is: James W Griffith, PhD, Kimberly Kenton, MD, MS, Christina Lewicky-Gaupp, MD, Todd Parrish, PhD, Jennie Yufen Chen, PhD, Margaret Mueller, MD; Study Coordinators: Sarah Buono, Maria Corona, Beatriz Menendez, Alexis Siurek, Meera Tavathia, Veronica Venezuela, Azra Muftic, Pooja Talaty, Jasmine Nero. Dr. Helfand, Ms. Talaty, and Ms. Nero are at NorthShore University HealthSystem.

University of Michigan Health System, Ann Arbor, MI (DK099932): PI: J Quentin Clemens, MD, FACS, MSCI; Co-Is: Mitch Berger, MD, PhD, John DeLancey, MD, Dee Fenner, MD, Rick Harris, MD, Steve Harte, PhD, Anne P. Cameron, MD, John Wei, MD; Study Coordinators: Morgen Barroso, Linda Drnek, Greg Mowatt, Julie Tumbarello

University of Washington, Seattle Washington (DK100011): PI: Claire Yang, MD; Co-I: John L. Gore, MD, MS; Study Coordinators: Alice Liu, MPH, Brenda Vicars, RN

Washington University in St. Louis, St. Louis Missouri (DK100017): PI: Gerald L. Andriole, MD, H. Henry Lai; Co-I: Joshua Shimony, MD, PhD; Study Coordinators: Susan Mueller, RN, BSN, Heather Wilson, LPN, Deborah Ksiazek, BS, Aleksandra Klim, RN, MHS, CCRC

National Institute of Diabetes and Digestive and Kidney Diseases, Division of Kidney, Urology, and Hematology, Bethesda, MD: Project Scientist: Ziya Kirkali MD; Project Officer: Christopher Mullins PhD; NIH Personnel: Tamara Bavendam, MD, Robert Star, MD, Jenna Norton

Arbor Research Collaborative for Health, Data Coordinating Center (DK097776 and DK099879): PI: Robert Merion, MD, FACS; Co-Is: Victor Andreev, PhD, DSc, Brenda Gillespie, PhD, Gang Liu, PhD, Abigail Smith, PhD; Project Manager: Melissa Fava, MPA, PMP; Clinical Study Process Manager: Peg Hill-Callahan, BS, LSW; Clinical Monitor: Timothy Buck, BS, CCRP; Research Analysts: Margaret Helmuth, MA, Jon Wiseman, MS; Project Associate: Julieanne Lock, MLitt

Abstract:

Aims: The aims of this study were to assess completeness of voiding diaries in a research context, and to correlate diary data with patient-reported questionnaires.

Methods: Men and women enrolled in the Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN) were given a 3-day voiding and fluid intake diary to fill-out. Diaries were assessed for completeness and intake-output imbalances. They were assigned to one of four categories based on percentage of missing data and fluid imbalance: no diary submitted, unusable (>40% missing void or intake volumes, or unphysiological fluid imbalance), usable but not complete, and complete.

Results: A total of 1,064 participants were enrolled and 85% (n=902) returned the bladder diary. Of the diaries returned, 94% (n=845) had data on 3 separate days, 87% (n=786) had no missing intake volumes, 61% (n=547) had no missing voided volumes, and 70% (n=635) had a fluid imbalance within 3L across the three-day time period, resulting in 50% (n=448) of participants with 100% complete diaries. Younger age was associated with a higher likelihood of not submitting a diary, or submitting an unusable diary. Women had a higher likelihood of submitting an unusable diary, or a usable but incomplete diary.

Conclusion: Overall, 50% of LURN participants returned voiding diaries with perfectly complete data. Incomplete data for voided volumes was the most common deficiency. There was only moderate correlation between diary data and questionnaire responses, indicating that diaries are a source of unique information.

Introduction:

Bladder diaries are a rich source of relatively objective information on the voiding and fluid consumption habits of patients with lower urinary tract symptoms (LUTS). They are useful in identifying potential causes of LUTS, guiding behavior modifications, and assessing treatment outcomes. However, the usefulness of bladder diaries is largely dependent upon complete and accurate data. Diaries require significant patient effort to complete, which can result in missing data[1] .

Several studies have assessed the utility of bladder diaries but most were limited by inclusion of a single sex [2–7], a single disease process, or a small study population [1,2,11,3–10]. Self-reported measures of urinary symptoms are much faster to complete, but it is also not known whether these are sufficient to establish a clear clinical picture or if the diary is also needed [12]. Some studies that have compared bladder diaries and symptom scores derived from self-report items suggest that the bladder diary is superfluous at least for overactive bladder [5] and incontinence[7] in women , but this needs to be assessed in a larger sex-mixed population with different LUTS. The primary aim of this study was to describe the completeness and accuracy of a three-day bladder diary in a large cohort of men and women seeking care for LUTS. Our hypotheses were that (1) most patients can accurately capture voiding habits and fluid intake in a diary, (2) voiding diary data and questionnaire data will correlate highly for daytime symptoms, and (3) voiding diary data and questionnaire data will correlate poorly for nighttime symptoms.

Materials and Methods:

Study design and population

The Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN) recruited men and women from six U.S. tertiary care sites as part of a one-year prospective observational study. The study design and recruitment criteria have been previously detailed[13]. Participants were at least 18 years of age and presented to a LURN physician with at least one LUTS as assessed by the LUTS Tool [14] using a 1-month recall period. Participants with neurogenic bladder, major psychiatric disorders or other medical issues that might interfere with study participation (e.g., dementia, psychosis, etc.), as well as those who had difficulty reading or communicating in English, were excluded from the study. Data collected at a baseline visit included demographics, medical history, the Functional Comorbidity Index[15](FCI), body mass index (BMI) , a three-day bladder diary, patient-reported urinary symptoms, psychological symptoms, and quality of life.

Measures:

Bladder diaries in the LURN Observational Cohort Study were collected using a modified International Consultation on Incontinence Questionnaire (ICIQ) bladder diary [8], a validated diary with detailed information on participant intake and output over a three-day period. Study participants were instructed to enter the void and intake amount, intake type, the time each void and intake took place, and record their bladder sensation at the time of voiding (e.g. none, normal, urgency), pad changes, and incontinence episodes (“stress” or “urge” or “unknown”). Participants reported the times they went to bed and woke up for the day, allowing for the classification of daytime and nighttime events.

Diary quality was assessed using the following metrics: void and intake data on three separate days, proportion of voids and intakes with a missing volume, and fluid imbalance over the three-day period (calculated as total volume consumed minus total volume voided). Diaries with at least one void and intake recorded on three separate days and less than 40% missing void and intake volumes were classified as "usable", while any returned diaries that did not meet these criteria were classified as "unusable". Among the "usable" diaries, those with no missing volume data and a physiologically feasible fluid imbalance (<3L across all three days) were considered "complete," while others were defined as "usable, but not complete." Hence all complete diaries had to be a full three days long, have all voided volumes and intake volumes recorded accurately with no checkmarks or simple statements such as "a lot" or "little" and be fluid balanced on average over the three days with no more than 1 liter per day of imbalance indicative of missed recordings.

For the complete diaries, the following measures were derived (average per 24-hour period and total over the three 24-hour periods): number of voids, total voided volume, number of intakes, total intake volume, fluid balance, and number of incontinence episodes. Maximum voided volume and average volume per void were also calculated. Voids were split into day and night components for each 24-hour period, with day voids occurring during awake hours and night voids during sleeping hours. If waking and sleeping times were not indicated for a given 24-hour period, participant averages from other 24-hour periods were used. If waking and sleeping times were not indicated for a given diary, waking times of 6 AM and bed times of 12 AM were used for each of the

three 24-hour periods in that diary. These times were selected a priori as conservative estimates of typical waking and bed times.

Participants completed the LUTS Tool, a 44-item questionnaire assessing the frequency and bother of LUTS over the preceding week. LUTS Tool question #2, “During a typical day in the past week, how many times did you urinate during waking hours?” and question #3, “During a typical night in the past week, how many times did you wake up because you needed to urinate?” were used for self-reported measures of average number of day and night voids, respectively.

Statistical methods:

A multinomial logistic regression model was fitted to assess factors significantly associated with the probability of membership in each category of bladder diary completion (complete, usable, unusable, and not submitted). Candidate covariates included age, sex, race, BMI category, education level, employment status, FCI, and history of a psychiatric diagnosis. Metrics derived from the bladder diary were summarized with means, standard deviations, quartiles, and ranges. The association between bladder diary metrics was explored using r-squared and polychoric correlation coefficients were used to evaluate associations between metrics derived from the bladder diary and self-reported urinary symptoms. All statistical analyses were completed using SAS 9.4 (Cary, North Carolina).

Results:

The LURN Observational Cohort Study enrolled 1064 participants (519 men and 545 women) with a mean age of 58.8 ±14.1 SD years. Most participants (83%) were white, and the mean BMI was 30.1 ±6.9 kg/m^2 with 43% classified as obese (BMI > 30 kg/m^2). Participants had few comorbidities with a mean FCI of 2.3 ±2.0, including 19% with self-reported diabetes mellitus. About two-thirds of participants reported incontinence on the LUTS Tool. Details of the enrollment, patient demographics and clinical symptoms were reported previously[13] .

Among the participants, 902 (84.8%) returned a baseline bladder diary. Diaries were excluded for the following reasons (not mutually exclusive): having fewer than three days completed (n=57), failing to record an amount for more than 40% of voiding episodes (n=48) or intake volumes (n=4). This yielded 796 usable diaries (88%). When stringent criteria were applied, 49.7% (448 of 902 diaries returned) had complete volume data for all recorded intake and voiding events and had negligible fluid balance differences (i.e., <3L across all three days). Among those excluded from this category, the most common reason was missing a voided volume (n=258) (Figure 1).

Younger age was associated with a higher likelihood of not submitting a diary, or submitting an unusable diary. Female sex was associated with a higher likelihood of submitting an unusable diary, or a usable but incomplete diary. Lower education level was associated with a higher likelihood of submitting an unusable diary (Table 1).

Among the complete diaries, participants reported a mean of 8.8 ±3.2 voids per 24-hour period with a mean total voided volume of 1769 ±701 ml. They reported an average of

1.1 ± 2.0 incontinence episodes per 24-hour period. In terms of intake, participants consumed an average of 1812 ±681 mL per 24-hour period with an average of 6.3 ±2.2 separate intakes. This yielded an average fluid imbalance of -43 ±458 mL per 24-hour period (Table 2). The distribution of average 24-hour fluid imbalance among all complete diaries is shown in Figure 2, and complete descriptive statistics on voiding and intake information is available in Supplemental Table S1.

When assessing the utility of voiding diaries for guiding behavioral modification, 51% of participants with complete diaries had an average 24-hour number of voids exceeding 8, the typical threshold for defining urinary frequency [16] (Figure 3). About a third of participants reported an average 24-hour intake volume over 2L, and 93 (64%) of these participants had urinary frequency (≥8 voids/day). Among these 93 participants, 57 (61%) reported being at least "somewhat" bothered by their urinary frequency. There was a statistically significant association between average 24-hour number of voids and intake volume ($p < 0.001$); however, the R^2 value of 0.037 indicated that only 4% of the variance in voided volume was explained by intake volume (Table 3). 67% of participants consumed less than 2L per day suggesting that they were possibly fluid restricting.

Comparing the average number of day and night voids to self-reported day and nighttime frequency on the LUTS Tool showed statistically significant and positive polychoric correlation coefficients ($p < 0.001$). However, polychoric correlation coefficients were below 0.500 (0.454 for average day voids and 0.493 for average night voids) (Figure 4). Furthermore, most patients recorded values in their bladder diaries outside of the range indicated on the corresponding response to the LUTS Tool question (59% for average number of day voids and 69% for night voids) (Figure 5). Specifically, for average

number of day voids, 85% of patients answering “1-3 times”, 77% answering “8-10 times”, 83% answering “11-13 times”, and 95% answering “14 or more times” had diary values outside of the range in their response. Similarly, for average number of night voids, 56% of patients answering “1 time”, 79% answering “2 times”, 88% answering “3 times”, and 87% answering “4 or more times” had average diary values (rounded up to the nearest whole void) not matching these responses.

Discussion:

Although clinically helpful, bladder diaries are often viewed as burdensome to patients and study participants. We demonstrated that the vast majority (85%) of LURN participants were willing to fill-out a three-day fluid intake and voiding diary.

Despite the strict criteria used to classify a diary as “complete,” almost 50% of diaries met this threshold. An additional 39% yielded incomplete but usable data. This is not to say that the other diaries were uninterpretable (and may have had clinical utility), but for research purposes they were deemed not useful for this analysis. The most common missing data point was voided volume (i.e., participant recorded that they voided but did not measure the volume). This would be expected in a LUTS population where 66% have urgency [13] since it is difficult to measure urine output when rushing to void, or when urgency incontinence occurred prior to measuring. Also, it is not surprising that women struggled more with measuring voided volume since urgency incontinence is more common in women and urine collection in the seated position is more challenging than for men who can collect urine in the standing position. This highlights the need for a method to collect voided volume that is easier for participants to use, particularly women.

Few studies have addressed diary completeness, with most relying on test retest variability as a marker of completeness[2,8]. However, reliability only reflects similarities between days, and not necessarily a patient's ability (or willingness) to complete the diary. Two studies in LUTS patients reported on diary completeness, with "complete" diaries (based on various measures) in 53% (214/400) of men and women [8] and 57% of 110 women [2], similar to our results.

Analysis of voiding diary data collected from both men and women is not commonly performed[1,5,11] and no studies have objectively compared completeness of the collection. Ku et al. assessed the subjective burden of voiding diaries among 57 men and 105 women with LUTS on a two-, three-, or seven-day diary. Diary completeness was assessed by asking the patients ,"Do you believe that you recorded the diary accurately?" and, "How often did you omit to record?" and there was no difference between diary length, but these responses were not compared between sexes[1]. There was no difference in patient-reported burden between the sexes, based on age or education, but seven-day diaries were perceived as more burdensome compared to two or three days; however, based on the subjective assessment, the seven-day diary was not determined to have more omissions. In another small study of 21 men and 133 women with LUTS who completed a seven-day diary on two separate occasions, the test retest reliability was similar between men and women but no other assessment of diary completeness was performed[11]. Our approach to assessing quality of data collection did not rely on participant recollection of their ability to complete the diary, but on objective assessments of the data collected, which likely explains our low "complete" diary rate.

When the data from the bladder diary were compared to questionnaire responses, the polychoric correlations did not indicate a strong relationship between the two instruments for both day and night voids, and most patients had mismatching responses between the bladder diary and questionnaire for the two LUTS Tool questions considered. This was in contrast to our hypothesis that daytime frequency would be well captured by patient questionnaires. For both day and night frequency, there appears to be a tendency on questionnaire responses for patients at the low end of the spectrum to minimize the reported number and those at the higher end of the spectrum to exaggerate their symptoms. This is in contrast to other studies[3,5,11] that correlated symptom scores of nocturia and frequency to bladder diaries where good correlations were found on the frequency of voids reported as a mean; however, these studies combined all patient answers which includes those who under- and over-report. We have instead stratified those specific patients who are not reliable at reporting their urinary frequency that are at both ends of the spectrum.

A notable finding was that even with careful fluid intake and voids recorded many patients had significant fluid imbalances over the study period of three days (Figure 2). Given that these were community dwelling adults who had not started any new medications such as diuretics per protocol this is difficult to explain, but is a common clinical frustration when the volumes do not match. These imbalances are worthy of future investigation.

This analysis has several limitations. All participants volunteered for this study and received a monetary incentive to complete the voiding diary; we cannot conclude that similarly well-completed diaries could be produced from clinical patients. However, all

of these participants were also treatment-seeking patients and if patients were made aware of the clinical benefit of this exercise they would hopefully be as compliant. Also, we noted limited variability among diary variables across the three days and concluded that perhaps one day of collection is enough, but we did not compare participants collecting one day of data to those collecting all three since there may be a learning phenomenon. Also, our only method of detecting a completely missed (no checkmark) entry for a void or a fluid intake was fluid imbalance of over 1000ml on one day. If a patient missed a single void or intake this would likely not result in such a large imbalance and could be missed.

Conclusions:

Treatment-seeking participants in the LURN study were relatively successful at completing fluid intake and voiding diaries over three days with 50% (of 85% who returned a diary) completing near perfect diaries. The most difficult data for participants to record was voided volume, and this was most difficult for women, likely due to the difficulty in collecting voided urine. Urinary frequency and nocturia were not well captured on patient symptom scores with participants downplaying their frequency at the low end and exaggerating their frequency at the higher end when compared to the objective voiding diary.

References

1. Ku JH, Jeong IG, Lim DJ, Byun S. Voiding Diary for the Evaluation of Urinary Incontinence and Lower Urinary Tract Symptoms : Prospective Assessment of Patient Compliance and Burden. *Neurourol Urodyn*. 2004;335(June):331-335. doi:10.1002/nau.20027

2. Jimenez-cidre MA, Lopez-fando L, Esteban-fuertes M, et al. The 3-Day Bladder Diary Is a Feasible , Reliable and Valid Tool to Evaluate the Lower Urinary Tract Symptoms in Women. *Neurourol Urodyn*. 2015;34:128-132. doi:10.1002/nau

3. Song M, Chun J, Yoo DS, Han J, Choo M. Correlation of the Overactive Bladder Symptom Score , and the Voiding Diary and Urodynamic Parameters in Patients with Overactive Bladder Syndrome. *Low Urin Tract Symtpoms*. 2014;6(July 2013):180-184. doi:10.1111/luts.12035

4. Brummen HJ Van, Heintz APM, Vaart CH Van Der. The Association Between Overactive Bladder Symptoms and Objective Parameters From Bladder Diary and Filling Cystometry. *Neurourol Urodyn*. 2004;42(February 2003):38-42. doi:10.1002/nau.10162

5. Homma Y, Kakizaki H, Yamaguchi O, et al. Assessment of Overactive Bladder Symptoms: Comparison of 3-Day Bladder Diary and the Overactive Bladder Symptoms Score. *Urology*. 2011;77(1):60-64. doi:10.1016/j.urology.2010.06.044

6. Fitzgerald MP, Ayuste D, Brubaker L. How do urinary diaries of women with an overactive bladder differ from those of asymptomatic controls ? *BJU Int*. 2005;96:365-367. doi:10.1111/j.1464-410X.2005.05632.x

7. Bradley CS, Brown JS, Van Den Eeden SK, Schembri M, Ragins A, Thom DH. Urinary incontinence self-report questions: Reproducibility and agreement with bladder diary. *Int Urogynecol J*. 2011;22(12):1565-1571. doi:10.1007/s00192-011-1503-3

8. Bright E, Cotterill N, Drake M, Abrams P. Developing and Validating the International Consultation on Incontinence Questionnaire Bladder Diary. *Eur Urol*. 2014:1-7. doi:10.1016/j.eururo.2014.02.057

9. Groutz A, Blaivas JG, Chaikin DC, et al. NONINVASIVE OUTCOME MEASURES OF URINARY INCONTINENCE AND LOWER URINARY TRACT SYMPTOMS : A MULTICENTER STUDY OF MICTURITION DIARY AND PAD TESTS. *J Urol*. 2000;164(September):698-701.

10. Yap TL, Cromwell DC, Emberton M. A systematic review of the reliability of frequency-volume charts in urological research and its implications for the optimum chart duration. *BJU Int*. 2006;99:9-16. doi:10.1111/j.1464-410X.2006.06499.x

11. Brown JS, Naughton KSMC, Wyman JF, et al. Measurement Characteristics of a voiding diary for use by men and women with overactive bladder. *Urology*. 2003;61(4):802-809. doi:10.1016/S0090-4295(02)02505-0

12. Abdelmoteleb H, Kamel MI, Hashim H. The association between the ICIQ-LUTS

and the ICIQ-bladder diary in assessing LUTS. *Neurourol Urodyn*. 2017;36(6):1601-1606. doi:10.1002/nau.23162

13. 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*. 2017;199(4). doi:10.1016/j.juro.2017.10.035

14. Coyne KS, Barsdorf AI, Thompson C, et al. Moving towards a comprehensive assessment of lower urinary tract symptoms (LUTS). *Neurourol Urodyn*. 2012;31(4):448-454. doi:10.1002/nau.21202

15. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol*. 2005;58(6):595-602. doi:10.1016/j.jclinepi.2004.10.018

16. Ricci JA, Baggish JS, Hunt TL, et al. Coping strategies and health care - Seeking behavior in a US national sample of adults with symptoms suggestive of overactive bladder. *Clin Ther*. 2001;23(8):1245-1259. doi:10.1016/S0149-2918(01)80104-1

Tables

Table 1. Associated factors of having a complete diary (results of multinomial logistic regression)

	Not submitted vs. complete			*Unusable vs. complete*			*Usable but not complete vs. complete*		
Variable	***Odds ratio estimate***	***95% confidence interval***	***p-value***	***Odds ratio estimate***	***95% confidence interval***	***p-value***	***Odds ratio estimate***	***95% confidence interval***	***p-value***
Age (per 10 yr. increase)	0.739	[0.646, 0.845]	<0.001	0.738	[0.633, 0.859]	<0.001	0.933	[0.838, 1.038]	0.201
Sex (Female vs. male)	0.891	[0.605, 1.311]	0.557	2.188	[1.378, 3.475]	<0.001	1.844	[1.377, 2.468]	<0.001
Education (High school or less vs. more than high school)	2.337	[1.398, 3.904]	0.001	1.320	[0.663, 2.627]	0.430	1.166	[0.733, 1.854]	0.516

Table 2. Bladder diary metrics per 24 hour period (complete diaries only, n=448)

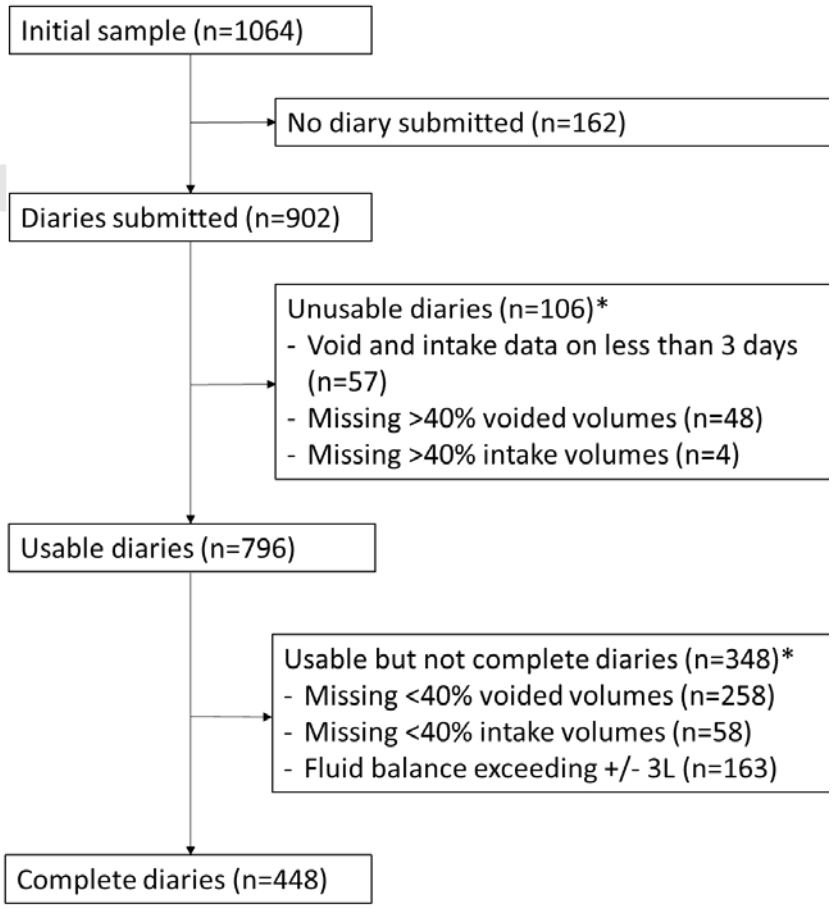
	Min.	*1st Quartile*	*Median*	*3rd Quartile*	*Max.*	*Mean*	*Std. Dev.*
Avg. number of voids	2.7	6.7	8.3	10.3	39.3	8.8	3.2
Total number of voids	8.0	20.0	25.0	31.0	118.0	26.3	9.5
Avg. voided volume (mL)	386.9	1237.2	1700.5	2139.2	4377.0	1768.6	701.3
Total voided volume (mL)	1160.8	3711.5	5101.5	6417.6	13130.9	5305.9	2103.9
Maximum voided volume (single void, mL)	118.3	300.0	414.0	591.5	1419.6	466.2	207.1
Avg. number of intakes	1.7	4.7	6.0	7.5	19.7	6.3	2.2
Total number of intakes	5.0	14.0	18.0	22.5	59.0	18.9	6.7
Avg. intake volume (mL)	522.5	1321.0	1725.2	2188.5	4495.2	1812.1	680.5
Total intake volume (mL)	1567.4	3962.9	5175.5	6565.4	13485.7	5436.2	2041.5
Avg. fluid balance (out minus in, mL)	-995.7	-386.9	-49.8	285.9	966.1	-43.4	458.4
Total fluid balance (out minus in, mL)	-2987.0	-1160.8	-149.3	857.6	2898.3	-130.3	1375.1
Avg. number of leaks	0.0	0.0	0.0	1.0	13.3	1.1	2.0
Total number of leaks	0.0	0.0	0.0	3.0	40.0	3.2	6.0

Table 3. Medians and interquartile ranges of bladder diary metric differences (complete diaries only, for each pairwise difference among the three 24-hr. blocks)

	24-hr. Block 1 – Block 2		*24-hr. Block 1 – Block 3*		*24-hr. Block 2 – Block 3*	
	Median	*Interquartile range*	*Median*	*Interquartile range*	*Median*	*Interquartile range*
Number of voids	0.0	-1.0 – 2.0	1.0	-1.0 – 3.0	1.0	-1.0 – 2.0
Voided volumes (mL)	-29.6	-407.0 – 354.9	236.6	-207.0 – 650.6	218.1	-147.9 – 709.8
Number of intakes	0.0	-1.0 – 2.0	1.0	-0.5 – 2.0	0.0	-1.0 – 2.0
Intake volumes (mL)	59.1	-321.4 – 436.2	147.9	-236.6 – 642.9	88.7	-244.0 – 532.3
Fluid balance (in minus out, mL)	-110.9	-532.3 – 421.4	29.6	-480.6 – 517.5	113.3	-412.8 – 621.1
Number of leaks	0.0	0.0 – 0.0	0.0	0.0 – 0.0	0.0	0.0 – 0.0

Figures

Figure 1. STROBE diagram of bladder diary completion



*Missing data and volumes are not mutually exclusive

Figure 2. Distribution of avg. 24-hr. fluid balance (output minus input, mL, complete diaries only, n=448)

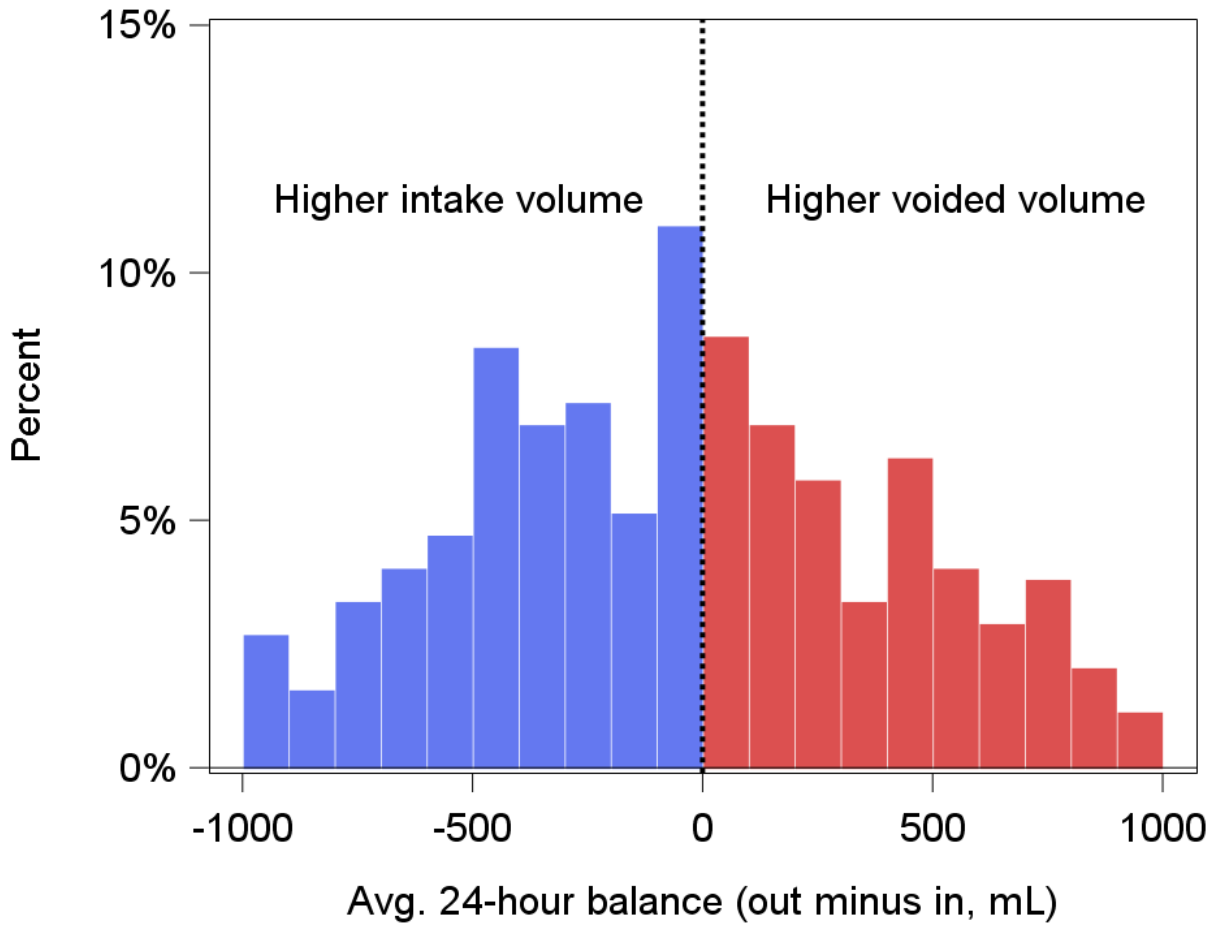


Figure 3. Scatterplot of avg. 24-hr. number of voids vs. intake volume

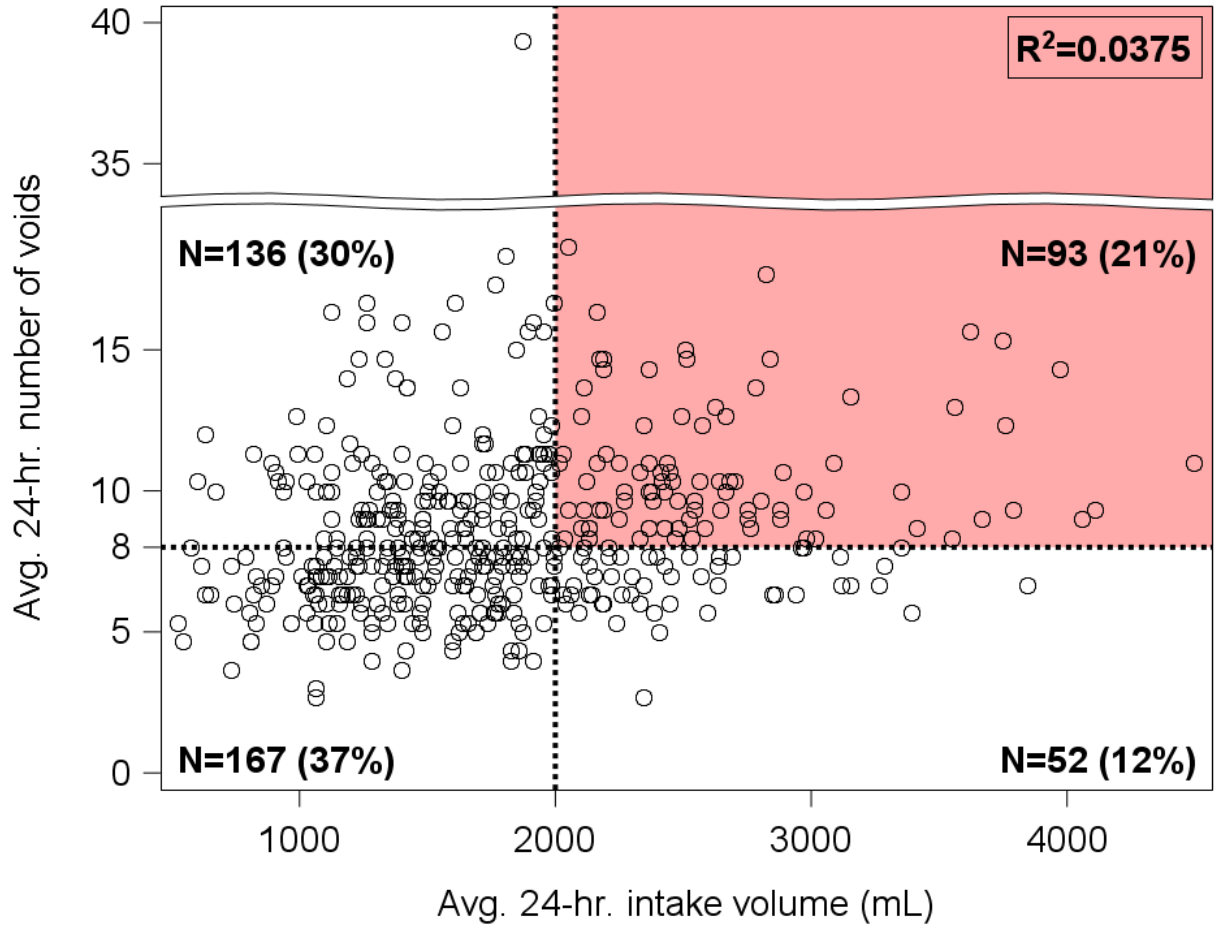


Figure 4. Distribution of bladder diary metrics within LUTS Tool response levels (complete diaries only)

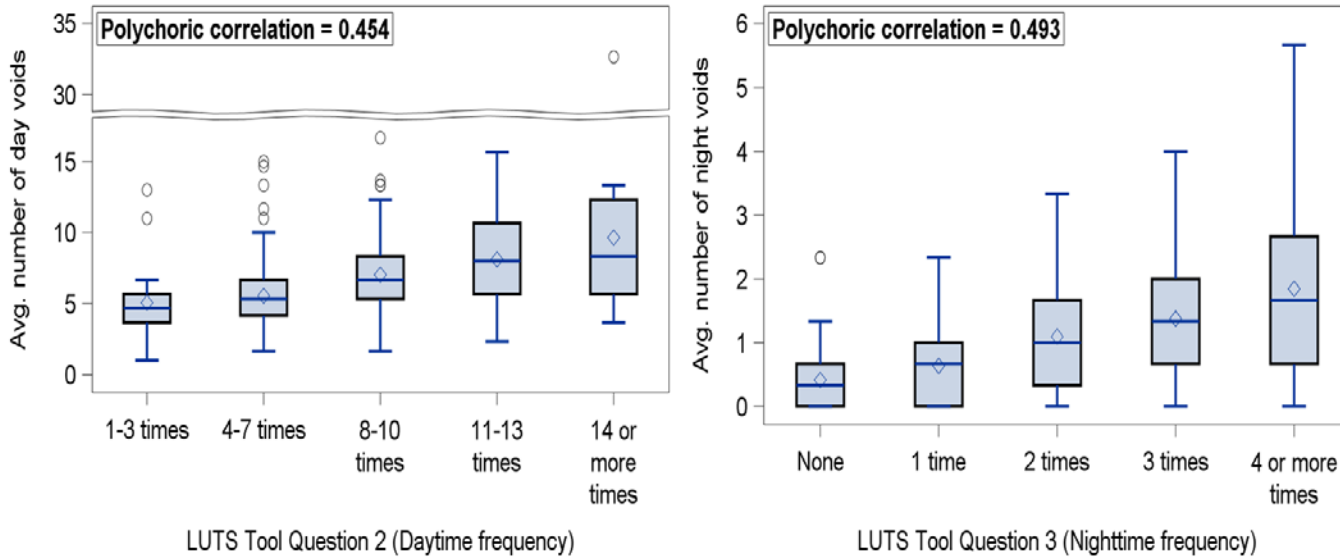


Figure 5. Comparison between bladder diary metrics and LUTS Tool responses (complete diaries only)

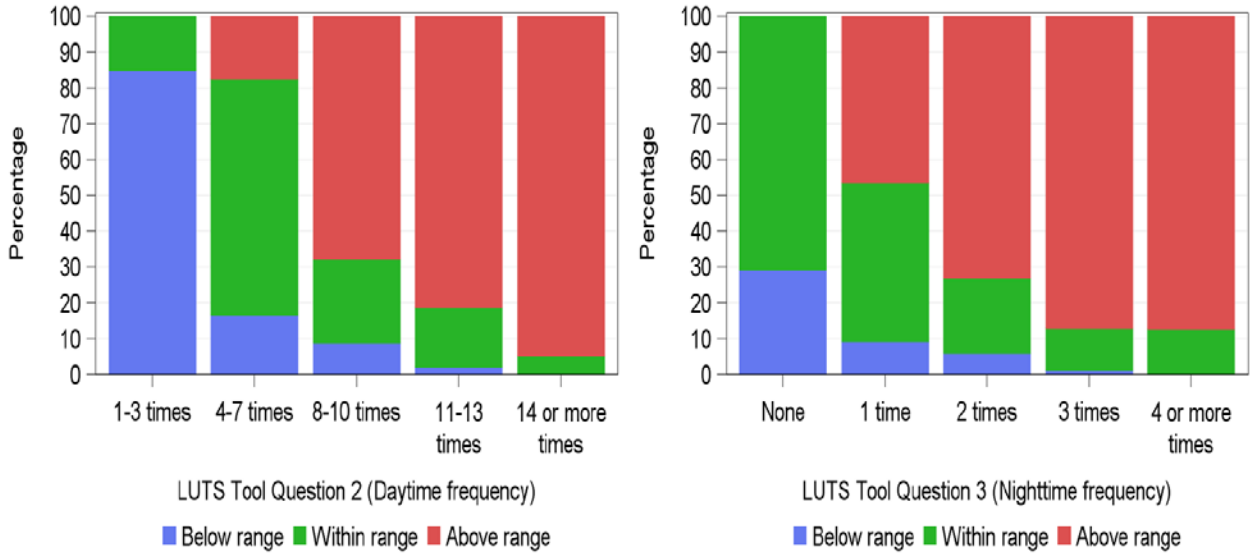


Figure legends

Figure 1. Strengthening the reporting of observational studies in epidemiology (STROBE) diagram of how participant's bladder diaries were categorized into "none submitted," "unusable," "usable," or "complete." Bladder diary groups and reasons for exclusion from groups are accompanied by relative sample sizes.

Figure 2. Distribution of average 24-hour fluid balance (mL), defined as average voided volume minus average intake volume, among complete diaries only (n=448). Blue bars represent values with higher average intake volume, and red bars represent values with higher average voided volume.

Figure 3. Scatterplot of average 24-hour number of voids vs. intake volume, among complete diaries only (n=448). Dotted line on the x- and y-axis represent typical thresholds for defining urinary frequency and excessive fluid intake, respectively. Shaded region represents participants with urinary frequency and excessive fluid intake.

Figure 4. Paneled graphic for distributions of a) average number of day voids by LUTS Tool question two response level, and b) average number of night voids by LUTS Tool question three response level. Diamonds within each box represent the mean for each response level and circles represent outliers. Lower, upper, and middle lines of the box represent the first quartile, median, and third quartile, respectively.

Figure 5. Paneled graphic for stacked bar charts of a) percentage of participants with LUTS tool question two responses higher, lower, or within range compared to bladder diary average number of day voids, and b) percentage of participants with LUTS tool question three responses higher, lower, or within range compared to bladder diary average number of night voids. Blue bars represent participants with LUTS tool responses lower than indicated on their bladder diary, green bars represent responses within range of their bladder diary, and red bars represent responses higher than their bladder diary metric. Average number of night voids were rounded to the nearest whole void.