

## **Use of the International Consultation on Incontinence Questionnaires bladder diary (ICIQ-BD) in men seeking therapy for lower urinary tract symptoms (LUTS)**

Hiroki Ito,<sup>1,2</sup> Paul Abrams,<sup>3</sup> Amanda L. Lewis,<sup>4,5</sup> Grace J. Young,<sup>4,5</sup> Peter S. Blair,<sup>4,5</sup> Nikki Cotterill,<sup>6</sup> J. Athene Lane<sup>4,5</sup>, Marcus J. Drake\*<sup>3,7</sup>

1 Department of urology, Yokosuka Kyosai Hospital, Yokosuka, Japan

2. Department of Urology, Yokohama City University Graduate School of Medicine, Yokohama, Japan

3. Bristol Urological Institute, North Bristol NHS Trust, Level 3, Learning and Research Building, Southmead Hospital, Bristol, BS10 5NB, UK

4. Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK

5. Bristol Randomised Trials Collaboration (BRTC), University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK

6. School of Health and Social Wellbeing, University of the West of England, Bristol, UK

7. Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

\*Marcus J. Drake (Corresponding author)

Email; [marcus.drake@bristol.ac.uk](mailto:marcus.drake@bristol.ac.uk)

Bristol Urological Institute, Level 3, Learning and Research Building, Southmead Hospital, Bristol, BS10 5NB UK, UK

Word Count; 2874

Key words; Lower urinary tract symptoms, nocturia, overactive bladder, bladder diary, questionnaire, urodynamics

## **Abstract**

*Context,* Completion rates and correspondence to other measures need to be established for the International Consultation on Incontinence Questionnaire (ICIQ) bladder diary (ICIQ-BD) in male lower urinary tract symptoms (LUTS) assessment.

*Objective,* To evaluate ICIQ-BD completion rates, frequency, volume and sensation reporting for men.

*Evidence Acquisition,* Evaluation of baseline data from the UPSTREAM randomised controlled trial evaluating 820 men at 26 UK hospitals, looking at ICIQ-BD, uroflowmetry, International Prostate Symptom Score (IPSS) and ICIQ-MLUTS.

*Evidence Synthesis,* 25.0% (205/820) provided complete voiding and bedtimes information for three days, 41.2% (338/820) omitted bedtimes information and the remainder omitted some or all voiding information. Median values (minimum - maximum) of 24 hour frequency, daytime frequency and night-time frequency were 9.7 (3.3 – 24.0), 7.7 (3.3 – 22.7) and 1.7 (0.0 – 5.7). Mean voided volume per micturition for daytime and night-time were 175.8ml  $\pm$  74.2 and 264.4  $\pm$  150.7 (P<0.001), respectively. For fully-completed diaries, daytime and night-time frequency showed weak-to-moderate correlation with symptom score questionnaires. More severe nocturia was generally reported in symptom scores than the ICIQ-BD. In patients with high bother for increased daytime frequency (symptom), the mean daytime frequency (ICIQ-BD) was 9.6  $\pm$  3.2, versus 7.6  $\pm$  2.2 for low bother (P<0.001). High bother for nocturia showed night-time frequency of 2.3  $\pm$  1.2, versus 1.5  $\pm$  1.1 for low bother (P<0.001). For fully- and partially-completed diaries, ICIQ-BD sensation scores correlated weakly with symptom scores. Voided volumes from the BD and uroflowmetry correlated weakly.

*Conclusion,* Two thirds of men (543/820) completed voiding information in the ICIQ-BD fully, but many omitted bedtimes information, limiting the ability to quantify nocturia and diagnose nocturnal polyuria.

*Patient Summary,* Most men with urinary symptoms complete a bladder diary fully but may fail to indicate bedtimes. Extra information from a diary helps support symptom questionnaires to explain a patient's urinary habits.

## **Introduction**

Lower urinary tract symptom (LUTS) are sub-divided into storage symptoms, voiding symptoms, and post-voiding symptoms. 80% of men aged over 40 experience at least one LUTS (1), which can compromise quality of life (QoL), employment and social activities (2). As the prevalence of LUTS increases with age, projected population ageing underlines the importance of effective management of male LUTS (3).

The bladder diary (BD) is an integral part of clinical LUTS assessment (4). It is used to assess the storage and voiding of patients with LUTS in a structured way, contributing information on potential factors relevant to initial management (5). Key information includes daytime and night-time frequencies, and voided volumes. Additional information can cover issues such as urgency, the pathognomonic symptom indicative of overactive bladder syndrome (OAB) (6). The BD supplements the symptom score assessment using validated patient reported outcome measures (PROs) (7), such as the International Prostate Symptom Score (IPSS) (8) and International Consultation on Incontinence Questionnaire (ICIQ) symptom score for Male LUTS (ICIQ-MLUTS) (9). The BD is essential for diagnosing nocturnal polyuria, which depends on identifying the urine volume made during the intended sleep period as a proportion of the 24-hour total. Hence, complete voiding information and times of going to bed and waking are all needed.

The ICIQ-BD is the only validated diary (6), and this study reports how well patients completed the diary in a research setting, evaluating completeness of submitted diaries (specifically relating to voided volumes and intended sleep period). Associations of symptom severity and symptom bother from questionnaires (IPSS and ICIQ-MLUTS) and ICIQ-BD were analysed for the baseline BD information from the UPSTREAM study. UPSTREAM (Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods) is a UK National Institute for Health Research funded multicenter pragmatic randomised controlled trial (ISRCTN56164274) looking at the diagnostic pathway for men being considered for prostate surgery to treat LUTS (9-13). The current study analysed both arms in conjunction, since only baseline UPSTREAM data was used.

## **Materials and methods**

The ICIQ-BD is a three day diary, specifically referring to three consecutive 24 hour periods, starting at 6 a.m. on the first day of recording, and concluding 72 hours later. It captures the times and volumes of all voids, bladder sensation for each void (using a BD-sensation scale), the times, types and volumes of liquid intake, and incontinence pad use. The patient is instructed to include "BED" and "WOKE" and times each day, so that intended sleep period ("night-time") can be distinguished from daytime. The IPSS (8) is a PRO with seven separate items scored from 0 to 5 (overall range 0 to 35), with higher scores indicating more severe symptoms. The ICIQ-MLUTS is a PRO comprising 13 questions, each scoring 0-4 in terms of severity, and the "bother" of each symptom (scored 0-10). For analysis, scores of 8-10 were used to denote substantial bother (9).

ICIQ-BD, IPSS, ICIQ-MLUTS and uroflowmetry data including maximum flow rate ( $Q_{max}$ ), voided volume

and post-void residual urine volume (PVR) were obtained from the baseline data of UPSTREAM. In brief, 820 men seeking further treatment for bothersome LUTS underwent standard LUTS assessment, and were randomised either to undergo or omit additional invasive urodynamic assessment. Inclusion criteria; men ( $\geq 18$  years) seeking further treatment for their bothersome LUTS. Exclusion criteria: unable to pass urine without a catheter; relevant neurological disease; active treatment for LUTS, on active surveillance for prostate or bladder cancer; previous prostate surgery; not medically fit, or unable to complete outcome assessments. The trial was approved by Oxford B Research Ethics Committee (reference: 14/SC/0237).

A paper copy of the ICIQ-BD was given to all participants at baseline (i.e. when enrolled in the trial), along with PROs (IPSS and ICIQ-MLUTS). Instructions on how to complete the ICIQ-BD are included on the first page of the diary. The BD-sensation scale is a five-point scale described in detail on the front page, with an abbreviated version on the back page. It includes two scores for normal voids, and three for voids with urgency or urgency incontinence (abbreviated scale; 0 - did not need to go, went just in case; 1 - normal desire to pass urine; 2 - had urgency but it passed away; 3 - had urgency but got to the toilet before leaking; 4 - had urgency and leaked). A BD-urgency frequency score (BDUFS) was calculated by adding the BD-sensation scale score of each void in a patient's urinary diary, and dividing this by the number of days recorded in the diary, analogous to the derivation of the Total Urgency-Frequency Score (TUFS) from the Patient Perception of Intensity of Urgency Scale (PPIUS) (14).

Completion rates for the bladder diaries by the 820 enrolled participants are reported as supplementary material in the baseline paper (13). Categorization in the current study used provision of voiding information for the full three days, and inclusion of intended sleep period information;

- Fully-completed BD: Three entire 24 hour periods of voiding information present, with documentation of "BED" and "WOKE" times for each of the intended sleep periods..
- Partially-complete BD: Three entire 24 hour periods of voiding information present, but missing one or more BED/ WOKE time.
- Non-complete BD: This was denoted if numerical voiding information was partly or entirely missing. For example, if the patient had placed a tick to indicate a non-measured void, the BD would be categorised as non-complete in the research context.

Clinical information derived from ICIQ-BD, IPSS, ICIQ-MLUTS and uroflowmetry at baseline were assessed according to the BD information. Correlations were analysed by Pearson's correlation coefficient. Comparison between two or three groups employed paired or unpaired t-tests or Tukey's test. Data were analysed with SPSS (SPSS 24, Chicago, Illinois). All statistical tests were two-sided and the strength of evidence against the null hypotheses were presented using p values ( $P=0.05$ ).

## **Results**

205/820 patients (25.0%) completed the ICIQ-BD fully over three days (complete voided volumes and intended sleep period information). 338/820 (41.2%) completed it partially (incomplete intended

sleep period information). The remaining 277/820 (33.8%) provided a non-complete ICIQ-BD (missing some numeric voiding data).

### **Clinical and demographic characteristics of completion groups**

The mean age of participants providing a non-complete ICIQ-BD ( $65.6 \pm 10.8$  years) was significantly younger than people who completed it fully ( $68.1 \pm 7.8$ ,  $P=0.009$ ) or partially ( $69.1 \pm 8.2$ ,  $P<0.001$ ) (Table 1). There was no difference between participants who completed the ICIQ-BD fully or partially, in terms of age or IPSS-QoL. Uroflowmetry findings, including  $Q_{max}$ , voided volume (VV) and PVR, were similar among the three groups. Mean total IPSS, ICIQ-MLUTS symptom severity score and bother score were not significantly different, except there was difference in the ICIQ-MLUTS symptom severity score between non-completers ( $18.9 \pm 7.4$ ) and partial completers ( $17.5 \pm 6.4$ ,  $P=0.039$ ). The mean IPSS-QoL of non-completers ( $4.3 \pm 1.4$ ) was significantly higher than for fully-completed ( $4.0 \pm 1.3$ ,  $P=0.041$ ) and partially-completed diaries ( $4.1 \pm 1.3$ ,  $P=0.036$ ).

### **Comparing fully- and partially- completed ICIQ-BD**

Mean voided volume per micturition (MVV) of fully-completed diaries ( $189.4 \pm 77.1$ ) was similar to the partially completed ones ( $177.9 \pm 77.2$ ,  $P=0.093$ ). Largest voided volume (LVV) over the 3 days of fully-completed diaries ( $339.2 \pm 131.4$ ) was significantly bigger than in partially completed ones ( $313.2 \pm 133.1$ ,  $P=0.024$ ). Mean BD-sensation scale score for fully- ( $1.6 \pm 0.6$ ) and partially-completed diaries ( $1.8 \pm 0.8$ ) were almost equivalent ( $P=0.249$ ). Likewise, mean BDUFS of fully- ( $16.4 \pm 9.4$ ) and partially-completed diaries were similar ( $15.4 \pm 10.2$ ,  $P=0.269$ ).

### **Comparison of BD with other PROs in men with fully-completed ICIQ-BD**

When evaluating how precisely matched night-time frequency was when derived from separate observation methods, there was considerable variability of 23-72% (IPSS and BD) and 25-83% (ICIQ-MLUTS and BD) (table 2). The IPSS and ICIQ-MLUTS reported more severe nocturia compared to ICIQ-BD for a high proportion of patients (Table 2). Median values (minimum - maximum) of 24-hour frequency, daytime frequency and night-time frequency from the ICIQ-BD were 9.7 (3.3 – 24.0), 7.7 (3.3 – 22.7) and 1.7 (0.0 – 5.7). MVV for the daytime and night-time were  $175.8 \pm 74.2$  and  $264.4 \pm 150.7$ , respectively ( $P<0.001$ ).

Correlation of recorded daytime and night-time frequency for the ICIQ-BD and the symptom score questionnaires is shown in Figures 1A-1D. Pearson's correlation coefficient between daytime frequency derived from the BD and the relevant questions in IPSS and ICIQ-MLUTS was 0.433 ( $P<0.001$ ) and 0.549 ( $P<0.001$ ) indicating weak-to-moderate correlation in both. Pearson's correlation coefficient between night-time frequency derived from BD and IPSS and ICIQ-MLUTS was 0.606 ( $P<0.001$ ) and 0.659 ( $P<0.001$ ) indicating moderate correlation in both.

Higher daytime and night-time frequency on the ICIQ-BD and the higher severity documented for the increased daytime frequency and nocturia questions in the symptom scores were similarly associated

with higher levels of bother. Pearson's correlation coefficient between daytime frequency derived from the BD and the related bother score in the ICIQ-MLUTS was 0.327 ( $P < 0.001$ ), indicating weak correlation. Pearson's correlation coefficient between night-time frequency derived from BD and the related bother score in the ICIQ-MLUTS was 0.450 ( $P < 0.001$ ), indicating weak correlation. Figure 2 shows the association between daytime and night-time frequency taken from the BD, and the data from the relevant ICIQ-MLUTS bother questions. Mean daytime frequency of patients with a bother score of  $\geq 8$  (out of a maximum 10) for increased daytime frequency was significantly higher than for bother scores of  $\leq 7$  ( $9.6 \pm 3.2$  vs.  $7.6 \pm 2.2$ ,  $P < 0.001$ ). Similarly, mean night-time frequency of patients with  $\geq 8$  bother score for nocturia was significantly higher than  $\leq 7$  ( $2.3 \pm 1.2$  vs.  $1.5 \pm 1.1$ ,  $P < 0.001$ ).

### **Pooled results for fully- and partially-completed ICIQ-BD**

205 fully-completed and 338 partially-completed diaries were also analysed jointly. Pearson's correlation coefficients between BD-sensation scale scores and the IPSS question related to urgency, ICIQ-MLUTS question for urgency severity and ICIQ-MLUTS question for urgency bother were 0.295 (Figure 3A,  $P < 0.001$ ), 0.342 (Figure 3B,  $P < 0.001$ ) and 0.316 ( $P < 0.001$ ) indicating weak correlation for each. Pearson's correlation coefficients between the BDUFS and the IPSS question for nocturia, ICIQ-MLUTS question for nocturia severity, and the ICIQ-MLUTS question for nocturia bother were 0.348 (Figure 3C,  $P < 0.001$ ), 0.345 (Figure 3D,  $P < 0.001$ ) and 0.350 ( $P < 0.001$ ) respectively, also indicating weak correlations.

Pearson's correlation coefficient between LVV of the BD and the uroflowmetry voided volume was 0.331 ( $P < 0.001$ , Figure 4A) indicating a relatively weak correlation. Pearson's correlation coefficient between the average VVs for the BD and uroflowmetry was 0.379 ( $P < 0.001$ , Figure 4B), also indicating weak correlation. Pearson's correlation coefficients between the MVV for the BD and the BD-urgency score and BDUFS were -0.116 ( $P = 0.007$ , Figure 4C) and -0.283 ( $P < 0.001$ , Figure 4D), respectively, indicating little correlation.

### **Discussion**

A fully-completed BD is taken as the most objective assessment of voiding frequency, and is specified in key guidance as an essential clinical evaluation (4, 15, 16). Validation of assessment tools is an important requirement to optimize assessment of LUTS, and the ICIQ-BD is currently the only diary to have been validated (17, 18). This study reports the findings for the ICIQ-BD in the context of a large research trial. The key information derived is as reported in the descriptive data listed in table 1, which summarises diary features divided into three groups according to voided volumes information (fully-, partially- or not completed). Two thirds of men provided complete information on voided volumes, with slightly younger patients with more severe LUTS being less likely to complete the voided volume information. The partially-completed diaries omitted some of the "BED" and "WOKE" times that are needed to divide the 24 hour period into daytime and the intended sleep period. For clinical evaluation or research in nocturia this is a problem. International Continence Society standardisations for defining

nocturia and nocturnal polyuria state that a BD should be evaluated (19, 20). Furthermore, nocturnal polyuria is important to identify for insight into underlying mechanisms of nocturia (21, 22). Hence, measures to improve the likelihood of including this information would be valuable in future versions of the ICIQ-BD. In our clinical experience with the ICIQ-BD, if the person has failed to enter the BED/WOKE times, it is usually possible to add them retrospectively, when going through the diary with the patient in the clinic/office. Hence, it is feasible to confirm or exclude nocturnal polyuria in most individuals. This is a vital step, as the diagnosis of nocturnal polyuria necessitates a different care pathway from that for other causes of nocturia, such as poor bladder emptying or overactive bladder (23). The clinical setting may also enable extrapolation of voided volume information to surmise missing volumes. For example, the ICIQ-BD instructs patients had placed a tick to indicate a non-measured void. For the current research analysis, missing volume information meant the diary was counted as non-completed.

In the ICS terminology standardisation (24), the symptom of increased daytime frequency (“complaint that voiding occurs more frequently during waking hours than previously deemed normal by the individual (or caregivers)”) must be distinguished from the sign of frequency (“number of voids during daytime (awake hours)”). The distinction is covered by including PROs to evaluate the symptom, and BDs to look at the sign. However, the overlap is considerable and higher daytime and night-time frequency on the ICIQ-BD were associated with higher levels of bother for both the increased daytime frequency and nocturia questions in the symptom scores in an almost proportionate manner.

The BD-sensation scale scores diverged from the symptom scores for urgency, indicating that this may be a separate interpretation of urgency. The BD also provides a feel for the potential fluctuation of the measured parameters over the course of the three day period. In contrast, the symptom score is likely to reflect the patient’s integration of their symptoms over time, which could either result from their wish to communicate the situation at its worst, or an attempt they have made to balance out the variance they may experience to reflect a middle ground between extremes. This may explain why symptom scores tended to report worse severity than was identified from the ICIQ-BD for daytime and night-time voiding frequency, as has been identified with other BDs (25). The BD should accurately reflect actual voiding frequency (day-time and nocturia) if it is completed contemporaneously. However, we did not evaluate whether men actually updated the BD soon after each void. As such, we cannot be certain whether the diary was a true reflection of voiding activity.

Similarities between the component items and also the scorings of both ICIQ-BD sensation scale and the BDUFS aligned with previous reports (26), suggesting that the BD-sensation scale performs in a similar manner to the PPIUS. The PPIUS scale has similar items to the ICIQ-BD sensation scale except that score 1 is described as “mild urgency” in PPIUS rather than “normal desire to pass urine”. We derived the BDUFS from the ICIQ-BD sensation scale in the same way that the TUFs is calculated from the PPIUS (26-28). The urgency scores in IPSS and ICIQ-MLUTS showed higher correlation with BDUFS than the BD- sensation scale score did. Storage LUTS are considered a relative contraindication for interventional therapy of voiding LUTS. Accordingly, further research could establish which urgency

marker (symptom scores, BD- sensation scale score or BDUFS) provides a better indicator of adverse outcome for interventional treatment for voiding LUTS.

The IPSS and ICIQ-MLUTS both showed considerable discrepancy to documented night-time frequency in the BD. The problem may reflect the unclear wording of the symptom score questions for nocturia, which do not make it clear that the last void at bedtime and the first void when rising in the morning do not count towards nocturia severity (19). Thus, clinical situations where nocturia or daytime frequency are a major therapy target potentially should include a BD as a main outcome measure, and not rely solely on symptom scores.

To see the relationship between ICIQ-BD information and uroflowmetry parameters, the VVs derived from both tests were compared. LVV and MVV derived from the BD showed weak correlation with VV derived from uroflowmetry. This suggests that it is appropriate to scrutinise the BD when interpreting free flowmetry data, since a large discrepancy between the BD volumes and the flowmetry VV would cast doubt on whether the uroflowmetry findings were representative of daily voiding experience. This could reflect that patients in flow clinics are generally asked to drink more than usual and to defer voiding as long as possible, in a way they are unlikely to do at home.

The fact that some patients did not fully complete their BD is a limitation which means that the conclusions may be biased by the selective nature of likelihood of full completion. The diaries were obtained in the setting of a clinical trial in secondary care, which may differ from use in routine clinical practice. The findings are derived from baseline data, and further research is needed to evaluate how they relate to therapy outcome. Further research could also evaluate which parameter is the best descriptor for OAB, since the BD sensation scale scores diverged from the symptom scores for urgency, indicating that they provide a separate interpretation.

### **Conclusions**

In a research trial for male LUTS, two thirds of men (543/820) completed the voiding information in the ICIQ-BD fully, but a high proportion omitted bedtimes information, limiting ability to interpret nocturia, and in particular to diagnose nocturnal polyuria. Partially-completed diaries provide additional information for clinical staff to interpret urinary frequency, urinary VV and urgency, alongside questionnaires and uroflowmetry data.

### **Financial disclosure**

The UPSTREAM project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme (project number 12/140/01). The views and opinions expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. This study was designed and delivered in collaboration with the BRTC, a UKCRC Registered Clinical Trials Unit (CTU), which, as part of the Bristol Trials Centre (BTC), is in receipt of NIHR CTU support funding. The UPSTREAM trial was sponsored by North Bristol NHS Trust (NBT), Bristol, UK. Study data were collected and managed using REDCap (29), hosted at the University of



Bristol.

The authors thank all the men who participated in the UPSTREAM trial. We also thank the participating NHS hospitals, in particular the principal investigators, research nurses, clinical trial assistants, administrators, and all other support staff, for their dedication and hard work.

### **Non-financial disclosure**

PA is Chairperson for the International Consultations on Incontinence. MJD and NC are Advisory Board members of the ICIQ research project.

### **References**

1. Irwin DE, Milsom I, Kopp Z, Abrams P, Artibani W, Herschorn S. Prevalence, severity, and symptom bother of lower urinary tract symptoms among men in the EPIC study: impact of overactive bladder. *European urology*. 2009;56(1):14-20.
2. Kim TH, Han DH, Ryu DS, Lee KS. The Impact of Lower Urinary Tract Symptoms on Quality of Life, Work Productivity, Depressive Symptoms, and Sexuality in Korean Men Aged 40 Years and Older: A Population-Based Survey. *International neurourology journal*. 2015;19(2):120-9.
3. Milsom I, Irwin DE. A Cross-Sectional, Population-Based, Multinational Study of the Prevalence of Overactive Bladder and Lower Urinary Tract Symptoms: Results from the EPIC Study. *European Urology Supplements*. 2007;6(1):4-9.
4. Gratzke C, Bachmann A, Descazeaud A, Drake MJ, Madersbacher S, Mamoulakis C, et al. EAU Guidelines on the Assessment of Non-neurogenic Male Lower Urinary Tract Symptoms including Benign Prostatic Obstruction. *Eur Urol*. 2015;67(6):1099-109.
5. Bright E, Drake MJ, Abrams P. Urinary diaries: evidence for the development and validation of diary content, format, and duration. *Neurourol Urodyn*. 2011;30(3):348-52.
6. Bright E, Cotterill N, Drake M, Abrams P. Developing and validating the International Consultation on Incontinence Questionnaire bladder diary. *European urology*. 2014;66(2):294-300.
7. Cameron AP, Wiseman JB, Smith AR, Merion RM, Gillespie BW, Bradley CS, et al. Are three-day voiding diaries feasible and reliable? Results from the Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN) cohort. *Neurourology and urodynamics*. 2019;38(8):2185-93.
8. Barry MJ, Fowler FJ, Jr., O'Leary MP, Bruskewitz RC, Holtgrewe HL, Mebust WK, et al. The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. *J Urol*. 1992;148(5):1549-57.
9. Ito H, Young GJ, Lewis AL, Blair PS, Cotterill N, Lane JA, et al. Grading Severity and Bother Using the International Prostate Symptom Score and International Consultation on Incontinence Questionnaire Male Lower Urinary Tract Symptoms Score in Men Seeking Lower Urinary Tract Symptoms Therapy. *J Urol*. 2020;204(5):1003-11.
10. Bailey K, Abrams P, Blair PS, Chapple C, Glazener C, Horwood J, et al. Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods (UPSTREAM) for diagnosis and

management of bladder outlet obstruction in men: study protocol for a randomised controlled trial. *Trials*. 2015;16:567.

11. Young GJ, Lewis AL, Lane JA, Winton HL, Drake MJ, Blair PS. Statistical analysis plan for the Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods (UPSTREAM). *Trials*. 2017;18(1):455.

12. Drake MJ, Lewis AL, Young GJ, Abrams P, Blair PS, Chapple C, et al. Diagnostic Assessment of Lower Urinary Tract Symptoms in Men Considering Prostate Surgery: A Noninferiority Randomised Controlled Trial of Urodynamics in 26 Hospitals. *Eur Urol*. 2020;78(5):701-10.

13. Lewis AL, Young GJ, Abrams P, Blair PS, Chapple C, Glazener CMA, et al. Clinical and Patient-reported Outcome Measures in Men Referred for Consideration of Surgery to Treat Lower Urinary Tract Symptoms: Baseline Results and Diagnostic Findings of the Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods (UPSTREAM). *Eur Urol Focus*. 2019;5(3):340-50.

14. Chapple CR, Drake MJ, Van Kerrebroeck P, Cardozo L, Drogendijk T, Klaver M, et al. Total urgency and frequency score as a measure of urgency and frequency in overactive bladder and storage lower urinary tract symptoms. *BJU international*. 2014;113(5):696-703.

15. Marshall SD, Raskolnikov D, Blaker MH, Hashim H, Kupelian V, Tikkinen KA, et al. Nocturia: Current Levels of Evidence and Recommendations From the International Consultation on Male Lower Urinary Tract Symptoms. *Urology*. 2015;85(6):1291-9.

16. Sakalis VI, Karavitakis M, Bedretdinova D, Bach T, Bosch J, Gacci M, et al. Medical Treatment of Nocturia in Men with Lower Urinary Tract Symptoms: Systematic Review by the European Association of Urology Guidelines Panel for Male Lower Urinary Tract Symptoms. *Eur Urol*. 2017;72(5):757-69.

17. Bright E, Cotterill N, Drake M, Abrams P. Developing a validated urinary diary: phase 1. *Neurourology and Urodynamics*. 2012;31(5):625-33.

18. Bright E, Cotterill N, Drake M, Abrams P. Developing and validating the International Consultation on Incontinence Questionnaire bladder diary. *European Urology*. 2014;66(2):294-300.

19. Hashim H, Blaker MH, Drake MJ, Djurhuus JC, Meijlink J, Morris V, et al. International Continence Society (ICS) report on the terminology for nocturia and nocturnal lower urinary tract function. *Neurourol Urodyn*. 2019;38(2):499-508.

20. Hofmeester I, Kollen BJ, Steffens MG, Bosch JL, Drake MJ, Weiss JP, et al. Impact of the International Continence Society (ICS) report on the standardisation of terminology in nocturia on the quality of reports on nocturia and nocturnal polyuria: a systematic review. *BJU Int*. 2015;115(4):520-36.

21. Weiss JP, Ruud Bosch JL, Drake M, Dmochowski RR, Hashim H, Hijaz A, et al. Nocturia think tank: Focus on nocturnal polyuria: ICI-RS 2011. *Neurourol Urodyn*. 2012;31(3):330-9.

22. Hofmeester I, Kollen BJ, Steffens MG, Bosch JL, Drake MJ, Weiss JP, et al. The association between nocturia and nocturnal polyuria in clinical and epidemiological studies: a systematic review and meta-analyses. *The Journal of Urology*. 2013.

23. Gulur DM, Drake MJ. Management of overactive bladder. *Nat Rev Urol*. 2010;7(10):572-82.
24. D'Ancona C, Haylen B, Oelke M, Abranches-Monteiro L, Arnold E, Goldman H, et al. The International Continence Society (ICS) report on the terminology for adult male lower urinary tract and pelvic floor symptoms and dysfunction. *Neurourol Urodyn*. 2019;38(2):433-77.
25. Yap TL, Cromwell DA, Brown C, van der Meulen J, Emberton M. The relationship between objective frequency-volume chart data and the I-PSS in men with lower urinary tract symptoms. *Eur Urol*. 2007;52(3):811-8.
26. Mathias SD, Crosby RD, Nazir J, Klaver M, Drogendijk T, Hakimi Z, et al. Validation of the Patient Perception of Intensity of Urgency Scale in patients with lower urinary tract symptoms associated with benign prostatic hyperplasia. *Value Health*. 2014;17(8):823-9.
27. Cartwright R, Srikrishna S, Cardozo L, Robinson D. Validity and reliability of the patient's perception of intensity of urgency scale in overactive bladder. *BJU international*. 2011;107(10):1612-7.
28. Notte SM, Marshall TS, Lee M, Hakimi Z, Odeyemi I, Chen WH, et al. Content validity and test-retest reliability of Patient Perception of Intensity of Urgency Scale (PPIUS) for overactive bladder. *BMC urology*. 2012;12:26.
29. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-81.

### Table legends

Table 1: Descriptive data and comparison for the ICIQ-bladder diaries according to completeness.

Table 2: Proportions of patients with matching values for night-time frequency on the ICIQ-BD and the symptom score questionnaires.

### Figure legends

Figure 1. Box plot of daytime and night-time frequency derived from IPSS (left) and ICIQ-MLUTS severity (middle) and ICIQ-MLUTS bother, plotted against frequency data derived from bladder diaries

Figure 2. (A) Histograms of daytime frequency and night-time frequency from the fully-completed bladder diaries according to ICIQ-MLUTS Q13 (related to daytime frequency) and Q14 (night-time frequency). Mean daytime frequency of patients with substantial bother (bother score of at least 8 in Q13) ( $9.6 \pm 3.2$ ) was significantly higher than for the rest of the cohort ( $7.6 \pm 2.2$ ,  $P < 0.001$ ). Similarly, mean night-time frequency of patients scoring  $\geq 8$  in Q14 ( $2.3 \pm 1.2$ ) was significantly higher than those scoring  $\leq 7$  in Q14 ( $1.5 \pm 1.1$ ,  $P < 0.001$ ).

Figure 3. Box plots of the BD- sensation scale score (A) and BD-urgency frequency score (BDUFS) (C) against IPSS Q4 (related to urgency). Box plot of the BD-urgency score (B) and BDUFS (D) against ICIQ-MLUTS Q7 (related to urgency severity).

Figure 4. Scatter plots of the largest (A) and mean (B) voided volumes derived from bladder diaries and voided volume of uroflowmetry. Scatter plots of the BD- sensation scale score (C) and BD-urgency frequency score (BDUFS) (D) and mean voided volume derived from bladder diary.