



Ito, H., Young, G., Lewis, A. L., Blair, P. S., Cotterill, N., Lane, A. A., Sakamaki, K., Drake, M. J., & Abrams, P. H. (2020). Grading severity and bother using the IPSS and ICIQ-MLUTS scores in men seeking lower urinary tract symptoms therapy. *The Journal of urology*.
<https://doi.org/10.1097/JU.0000000000001149>

Peer reviewed version

Link to published version (if available):
[10.1097/JU.0000000000001149](https://doi.org/10.1097/JU.0000000000001149)

[Link to publication record on the Bristol Research Portal](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via American Urological Association at <https://doi.org/10.1097/JU.0000000000001149> . Please refer to any applicable terms of use of the publisher.

University of Bristol – Bristol Research Portal

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/>

Author's Accepted Manuscript

Grading severity and bother using the IPSS and ICIQ-MLUTS scores in men seeking lower urinary tract symptoms therapy

Ito H, Young GJ, Lewis AL, Blair PS, Cotterill N, Lane JA, Sakamaki K, Drake MJ, Abrams P

DOI: [10.1097/JU.0000000000001149](https://doi.org/10.1097/JU.0000000000001149)

Reference: JU-20-344

To appear in: *The Journal of Urology*

Accepted Date: 15 May 2020

Please cite this article as: Ito H, Young GJ, Lewis AL, Blair PS, Cotterill N, Lane JA, Sakamaki K, Drake MJ, Abrams P, Grading severity and bother using the IPSS and ICIQ-MLUTS scores in men seeking lower urinary tract symptoms therapy, *The Journal of Urology*® (2019), doi: [10.1097/JU.0000000000001149](https://doi.org/10.1097/JU.0000000000001149).

DISCLAIMER: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our subscribers we are providing this early version of the article. The paper will be copy edited and typeset, and proof will be reviewed before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to The Journal pertain.

Grading severity and bother using the IPSS and ICIQ-MLUTS scores in men seeking lower urinary tract symptoms therapy

*Hiroki Ito*¹, *Grace J. Young*^{2,3}, *Amanda L. Lewis*^{2,3}, *Peter S. Blair*^{2,3*}, *Nikki Cotterill*¹, *J. Athene Lane*^{2,3}, *Kentaro Sakamaki*⁴, *Marcus J. Drake*^{1,5**}, *Paul Abrams*¹

1. Bristol Urological Institute, North Bristol NHS Trust, Level 3, Learning and Research Building, Southmead Hospital, Bristol, BS10 5NB UK
2. Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS UK
3. Bristol Randomised Trials Collaboration (BRTC), as part of the Bristol Trials Centres (BTC), University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS UK (*1-5 Whiteladies Rd, Clifton, Bristol, BS8 1NU)
4. Center for Data Science, Yokohama City University, Yokohama, Japan
5. Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Hiroki Ito. Email; pug_daikichi@yahoo.co.jp. North Bristol NHS Trust, Bristol Urological Institute, Level 3, Learning & Research Building, Southmead Hospital, Bristol, BS10 5NB, UK

Grace J. Young. Email; grace.young@bristol.ac.uk. Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK

Amanda L. Lewis. Email; amanda.lewis@bristol.ac.uk. Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK

Peter S. Blair. Email; P.S.Blair@bristol.ac.uk. Population Health Sciences, Bristol Medical School, University of Bristol, 1-5 Whiteladies Rd, Clifton, Bristol, BS8 1NU, UK

Nikki Cotterill. Email; nikki.cotterill@bui.ac.uk. North Bristol NHS Trust, Bristol Urological Institute, Level 3, Learning & Research Building, Southmead Hospital, Bristol, BS10 5NB, UK

J. Athene Lane. Email; athene.lane@bristol.ac.uk. Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK

Kentaro Sakamaki. Email; sakamaki@yokohama-cu.ac.jp. Center for Data Science, Yokohama City University, 22-2 Seto, Kanazawa-ku, Yokohama 236-0027, Japan.

Marcus J. Drake (**Corresponding author). Email; marcus.drake@bristol.ac.uk. Bristol Medical School, Learning & Research Bldg, Southmead Hospital, Bristol, BS10 5NB, UK

Paul Abrams. Email; paul.abrams@bui.ac.uk. North Bristol NHS Trust, Bristol Urological

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

Running head; IPSS and ICIQ-MLUTS in men seeking LUTS therapy

Key words; Lower urinary tract symptoms, questionnaire, urodynamics, IPSS, ICIQ-MLUTS

Grading severity and bother using the IPSS and ICIQ-MLUTS scores in men seeking lower urinary tract symptoms therapy

Abstract

Purpose;

To establish severity banding ranges, bother assessment and key item content in principal patient reported outcomes measures (PROMs) in men seeking therapy for lower urinary tract symptoms (LUTS).

Materials and Methods;

Data for International Prostate Symptom Score (IPSS) and International Consultation on Incontinence Questionnaire Male LUTS (ICIQ-MLUTS) were derived from a study evaluating 820 men at 26 UK hospitals. Receiver operating characteristic (ROC) curves were used to establish severity bandings.

Results;

Classification tree showed that thresholds between mild-moderate and moderate-severe severity bands were 15 and 27 respectively for IPSS, 16 and 26 for ICIQ-MLUTS/severity, and 22 and 81 for ICIQ-MLUTS/bother. Highest area under ROC and lowest Akaike's information criteria of univariate logistic regression indicated that ICIQ-MLUTS/bother was more related to global quality of life (QoL) than were IPSS and ICIQ-MLUTS/severity. The symptoms affecting IPSS-QoL were only fully identified by ICIQ-MLUTS, because two key symptoms (urinary incontinence and post-micturition dribble) are not measured by IPSS. ICIQ-MLUTS demonstrated that bother of some LUTS is disproportionate to severity, and that persisting high bother levels following surgery are more likely due to storage (18-25%) and post-voiding (18-28%) LUTS than voiding LUTS (5-13%). Symptom improvement after surgery is uncertain if baseline IPSS-QoL score was <3.

Conclusions;

The severity threshold scores were measured for the two key LUTS PROMs, and the results indicate suitable categories of symptom severity for use in men referred for urological care.

The ICIQ-MLUTS measures all the LUTS affecting QoL, and includes individual symptom bother scores.

Introduction;

Lower urinary tract symptoms (LUTS) are categorized as storage symptoms, voiding symptoms and post-voiding symptoms. 80% of men aged >40 years experience at least one LUTS ¹, which can, for some, compromise quality of life (QoL) and social activities ². Projected population ageing underlines the importance of effective management of male LUTS ³.

The International Prostate Symptom Score (IPSS) ⁴ is a widely-used LUTS patient reported outcome measure (PROM). Seven separate items are scored from 0 to 5 (overall range 0 to 35), higher scores indicating more severe symptoms. To categorize patients' symptoms, the banding of symptom severity was set in 1992 as 1-7 mild, 8-19 moderate, 20-35 severe, derived from generic groups involving 210 LUTS and 108 non-LUTS patients ⁴. These symptom severity bandings are widely-used in clinical studies, both for establishing inclusion/exclusion criteria and categorizing baseline characteristics and measuring outcomes of pharmaceutical and surgical treatment ⁵⁻⁷. Men with voiding symptoms, storage symptoms and no symptoms have average IPSS scores of 16.8, 14.6 and 8.5, respectively ⁸. Thus, many patients apparently fall into the moderate range of IPSS severity despite describing themselves as 'symptom-free'. Accordingly, specific severity banding values in the context of LUTS assessment and treatment in specialist urology will facilitate decision-making and outcome assessment in this key sub-population.

The International Continence Society (ICS) instigated the International Consultation on Incontinence Questionnaires Male LUTS (ICIQ-MLUTS) module for men with LUTS seeking care ⁹. ICIQ-MLUTS is a validated PROM, given an A-grade by the International Consultation

on Incontinence. It has 13 questions, each scoring 0-4 in terms of severity, which are summed to give the overall score. "Voiding" and "Incontinence" subscores (incorporating urgency urinary incontinence (UI) and post-micturition dribble (PMD)) can also be derived. Many of the ICIQ-MLUTS/severity questions tie in with components of the IPSS, with good correspondence¹⁰, while the more comprehensive nature of the ICIQ-MLUTS includes symptoms omitted by IPSS which are potentially important to men, notably UI and PMD. Neither severity nor bother bandings for ICIQ-MLUTS have been established in any population. The "bother" of each symptom (scored 0-10) is identified separately in ICIQ-MLUTS, unlike the global QoL score used with the IPSS.

The aims of the current study were to use a healthcare-seeking male LUTS population: to establish severity banding ranges for ICIQ-MLUTS; to evaluate IPSS severity bands; to determine suitability of IPSS, ICIQ-MLUTS/severity or ICIQ-MLUTS/bother for assessing male LUTS; to evaluate the impact on QoL of key symptoms omitted by IPSS; and to establish the utility of individual item bother scores alongside a global QoL score for treatment planning.

Materials and methods

IPSS and ICIQ-MLUTS data were obtained from a two-arm randomised controlled trial set in 26 UK Hospitals (Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods (UPSTREAM); for full protocol, see ¹¹). In brief, 820 men seeking further treatment for bothersome LUTS underwent standard LUTS assessment and were randomised either to undergo or omit invasive urodynamic assessment. Inclusion criterion; men seeking further treatment for their bothersome LUTS. The exclusion criteria included inability to pass urine without a catheter; relevant neurological disease; active treatment or on active surveillance for prostate or bladder cancer; previous prostate surgery; not medically fit,

or unable to complete outcome assessments. The study was approved by Oxford B Research Ethics Committee (reference: 14/SC/0237). The current study is a *post-hoc* analysis (for statistical analysis plan see ¹⁴), evaluating men irrespective of the trial arm in which they participated.

IPSS and IPSS-QoL, ICIQ-MLUTS (13 questions, each scored 0-4 for severity and 0-10 for bother), urine flow rate and patient characteristics were measured before the start of LUTS treatment. The analysis set comprised participants with no missing data in all baseline IPSS and IPSS-QoL, ICIQ-MLUTS (severity and bother scores) questionnaires.

Receiver operating characteristics (ROC) curves and a univariate logistic regression model were used to evaluate the relationships of IPSS-QoL (\geq using 5 as the threshold value for bad QoL based on clinical criteria) with IPSS, ICIQ-MLUTS (severity and bother scores), age and uroflowmetry. The Akaike information criterion (AIC) is an estimator of prediction error, allowing estimation of the quality of each model ¹²; a low AIC value indicates a higher predictive ability for a given parameter. Thresholds of IPSS and ICIQ-MLUTS for severity bandings (binarization) were calculated based on examination of the Youden-index (rating index for a dichotomous diagnostic test) ¹³ from the ROC curves. Classification trees¹⁴ were also used for determination of thresholds for severity and bother bandings (three categories).

To evaluate the principal symptoms influencing QoL, the rankings of symptoms according to symptom severity scores in IPSS and ICIQ-MLUTS and bother score in ICIQ-MLUTS were summarized using the overall cohort of participants completing each of the questionnaires.

To investigate whether the IPSS-QoL and individual bother scores at baseline (before surgery) can predict the surgical treatment outcomes, surgical treatment outcomes measured by change of QoL improvement on IPSS-QoL and individual bother scores derived from ICIQ-MLUTS/bother were compared among patients.

Data were analyzed using R 3.6.1, pROC package ¹⁵ and rpart package ¹⁶.

Results

Patient selection and backgrounds

The analysis set comprised 672 participants who completed all questionnaires, including IPSS-QoL, IPSS, ICIQ-MLUTS/severity and ICIQ-MLUTS/bother (Fig. 1). The average age was 67.4 ± 9.0 years old (Table 1. For baseline descriptors of full trial population, see ¹⁰). Median (minimum - maximum) of the IPSS-QoL was 4 (0 - 6), IPSS was 19 (1 - 35), ICIQ-MLUTS/severity was 18 (3 - 40) and ICIQ-MLUTS/bother was 51 (0-130), respectively. Median values (minimum - maximum) of maximum flow rate (Q_{max}), voided volume (VV) and post-void residual (PVR) were 11.0 (1.4 – 58.3), 224.0 (7.0 – 947.0) and 100.0 (0.0 – 1600.0). The data of the overall study population was published previously ¹⁰.

New banding of IPSS, ICIQ-MLUTS symptom severity and bother scores

Univariate logistic regression analysis showed that IPSS and ICIQ-MLUTS/severity and ICIQ-MLUTS/bother are significant predictors of QoL ($P < 0.001$). The VV of uroflowmetry is also a significant QoL predictor ($P = 0.021$) (Table 2). The other objective values, including age and urine flow rate, did not show significant relationship with QoL (Table 2). Akaike's information criteria (AIC) of ICIQ-MLUTS/bother (732.015) was smaller than those of IPSS (762.002) and ICIQ-MLUTS/severity (762.99), with reliable coefficients for both continuous and categorized values (binary or 3-way categorization, as described below) (Supplementary Table S1). The areas under the ROC (AUROC) for IPSS, ICIQ-MLUTS/severity and ICIQ-MLUTS/bother were 0.728, 0.721 and 0.755, respectively (Figure 2).

For the widely-employed 3-way categorization (mild/moderate/severe), the classification tree showed 15 and 27 were the cut-off points between the mild-moderate and moderate-severe bandings of IPSS as a predictor of each patient's QoL impact (IPSS-QoL < 5 vs. ≥ 5) (Figure 3).

The classification tree for ICIQ-MLUTS/severity indicated 16 and 26 were the thresholds predicting QoL and that for ICIQ-MLUTS/bother indicated 22 and 81 (Figure 3). Univariate logistic regression analysis showed that the derived bandings (either binarization or 3 categories) of IPSS and ICIQ-MLUTS are significant predictors of QoL ($P < 0.001$). AIC of ICIQ-MLUTS/bother was smaller than those of IPSS and ICIQ-MLUTS/severity indicating ICIQ-MLUTS/bother can predict the global QoL more precisely than IPSS and ICIQ-MLUTS/severity.

Histograms of the distribution of patients within the severity bands of IPSS are presented in Figure 4. The proposed new IPSS banding (1-14 mild, 15-26 moderate, 27-35 severe) was able to categorize LUTS patients more consistently and reliably than the current banding (1-7 mild, 8-19 moderate, 20-35 severe).

Assessment of bother of individual LUTS

Figure 5 shows the distributions of scores for separate LUTS in each symptom scale. In ICIQ-MLUTS/severity scores, important components in the incontinence subscore were PMD (66.5%) followed by urgency UI (63.8%), leak for no reason (30.4%), stress UI (22.7%) and leak during sleep (15.2%). In total, these affected 81.6% (669/820 patients), falling to 68.2% (559/820 patients) if PMD was excluded.

Rankings of severe symptoms in IPSS and ICIQ-MLUTS were different (Table 3). The most severe symptom in IPSS ("More than half" and "Almost always") was increased daytime frequency (55%), followed by weak stream (47.9%) and incomplete emptying (38.3%) (Table 3). In ICIQ-MLUTS ("Most of the time" and "All of the time") rankings were "reduced strength (50%)" followed by "nocturia (47.8%)" and "incomplete emptying (36.5%)" (Table 3). For symptom bother, high bother scores (8, 9 and 10) in ICIQ-MLUTS identified nocturia (48.1%),

urgency (45.7%), urgency UI (39.2%), increased daytime frequency (35.5%), incomplete emptying (35.0%) and PMD (33.3%) as the most bothersome (Table 3).

Impact of assessment of bother of individual LUTS on prostate surgical treatment outcome

Of the 820 patients, 291 (35.5%) underwent relevant surgery. They experienced bigger change of IPSS-QoL score ($- 2.44 \pm 1.92$) than non-surgically treated patients ($- 0.68 \pm 1.49$) ($P < 0.001$) (Supplementary Figure S1).

The bother associated with LUTS following surgery for the men who completed each scores at baseline and after treatment, and also underwent surgery during the trial, was evaluated. The IPSS-QoL global scores (Supplementary Table S2A) for men with a baseline of 5-6 improved (i.e. fell by at least 1) in 89% and were unchanged in 9-11% after surgery. For baseline scores of 3-4, improvement was identified in 78-83%, and worsening in 10-13%. For baseline score of 2 (only 14 men), there was 50% chance of improvement and 43% worsened. Looking at bother associated with individual LUTS measured with the ICIQ-MLUTS (Table S2B), men with a high bother score of 8-10/10 at baseline were more likely to have persisting high bother scores after surgery for storage symptoms (nocturia 21.8%, urgency/urgency UI 17.5/24.7%, increased frequency 25%) and post-voiding symptoms (PMD 27.8%, sensation of incomplete emptying 18.2%) than for voiding symptoms (slow stream 5.4%, straining 13.2%).

Discussion

Evaluation of the severity of LUTS is crucial for making treatment recommendations and for evaluation of treatment outcome. Questionnaires indicate the status of LUTS and each patient's QoL. However, their continuous values are often widely spread when viewed across

an entire population, making it difficult to place the severity of LUTS for an individual reliably in comparison with the overall cohort. To counteract this issue, IPSS severity bandings have been widely used as an inclusion criterion in drug trials and guidelines. However, the utility of the bandings in the assessment of male LUTS was previously unknown. No prior study reassessed severity bandings of the IPSS or determined severity banding for the ICIQ-MLUTS, as far as we are aware. Therefore, we developed new banding ranges of IPSS and ICIQ-MLUTS in a large study of men referred for treatment of LUTS. The study benefitted from a large cohort of patients studied with extensive resources, which allowed detailed assessment and follow up.

Proposed new banding of IPSS, and ICIQ-MLUTS symptom severity and bother scores

The classification trees were utilized to determine the possible severity thresholds. Lower Akaike's information criteria (AIC) for both IPSS and ICIQ-MLUTS demonstrated that the questionnaires have a bigger impact and more important role in quantifying patients' QoL than other tests, notably uroflowmetry values. This is in accordance with previous studies indicating that the relationship between symptoms and uroflow variables is inconsistent¹⁷.

The present study confirms that the severity of a symptom does not necessarily anticipate how bothersome it is. The study newly introduces the appropriate threshold values for the moderate and severe categories of ICIQ-MLUTS, and re-evaluates them for the IPSS in the key population of men seeking urological care for LUTS. The thresholds were identified as: 15 and 27 for IPSS; 16 and 26 for ICIQ-MLUTS/severity; and 22 and 81 for ICIQ-MLUTS/bother. The current thresholds for IPSS (8 and 20) which are generally used, in both clinical practice and research, appear to classify LUTS patients inappropriately for this population, potentially yielding too many false positives. The principal reason is explained by the fact that the original banding was developed with generic groups involving LUTS and

“non-LUTS patients” (approximately half of whom had moderate IPSS scores) ⁴. However, for urologists and other clinicians, the essential role of these questionnaires is in evaluation of patients’ LUTS and QoL, rather than discriminating between LUTS and non-LUTS men. Univariate analysis also showed independent and significant relationships of the derived bandings of IPSS and ICIQ-MLUTS to the QoL. Accordingly, IPSS and ICIQ-MLUTS can predict global QoL both as a continuous value, and also as a categorical value when categorized by the proposed new bandings.

Impact of assessment of bother of individual LUTS, especially storage symptoms, on patients’ QoL and prostate surgical treatment outcome

Several clinical studies have investigated nocturia ^{18, 19}, increased daytime frequency ²⁰ and urgency (overactive bladder) ¹, which are thus already recognized as substantially bothersome ²¹⁻²³. A few previous studies compared bother associated with individual LUTS, and studies showed that the most bothersome symptoms in men are increased daytime frequency ²³, and urgency and urgency UI symptoms ²². The most bothersome individual symptoms were different for bother scores of ICIQ-MLUTS (nocturia) and symptom severity scores from ICIQ-MLUTS (increased daytime frequency) and IPSS (weak stream). The inclusion of nocturia seems relevant from clinical experience, while the absence of incontinence evaluation from the IPSS is problematic. Interpreting the AUROC and AIC analyses indicates that the ICIQ-MLUTS/bother scores facilitate understanding global QoL in male LUTS patients, and suggests that separate item bother scoring is more relevant than symptom severity for understanding QoL. ICIQ-MLUTS/bother can predict the global QoL more precisely than IPSS and ICIQ-MLUTS/severity, and identifies individual item contributions. In clinical practice, this means that appreciation of the bother scores for each symptom will enable direction of treatment towards the most bothersome symptom(s), with a

greater chance of benefit to overall QoL.

Recommendation of ICIQ-MLUTS vs IPSS in male LUTS assessment

The ICIQ-MLUTS includes incontinence and PMD, which significantly affect QoL and had more than 60% prevalence. With the diversity of LUTS patients²⁴, in both clinical and research settings, the individual symptom bother scores of ICIQ-MLUTS help discern the problematic symptom(s) responsible for the overall QoL impact, and thus identify the principle therapy need(s). Thus, ICIQ-MLUTS provides a more comprehensive insight, which exceeds the performance of IPSS, in the diagnostic assessment pathway.

Limitations

There are a few limitations in this study. Some patients did not fully complete the questionnaires, so they had to be excluded from the data sets. The cohort of this study consisted of men seeking therapy for LUTS, and it is in this population specifically that the findings should be interpreted. The approach taken employs a single method (using IPSS and ICIQ to predict dichotomized IPSS-QoL) to establish severity cut-offs. Reviewing the same question with different methods would be appropriate to consolidate the conclusions. Furthermore, a prospectively-derived dataset would be useful to confirm the current *post-hoc* analysis. This should include patients with isolated storage, voiding or post-voiding LUTS, or a mixed LUTS presentation, aiming to identify responses to specific therapy.

Conclusions

In a male LUTS population seeking further treatment, including the possibility of prostate surgery, the banding of symptom severity and bother was established for ICIQ-MLUTS, with

thresholds to denote boundaries between mild-moderate and moderate-severe of 16 and 26 respectively for severity and 22 and 81 for bother. The currently employed equivalent thresholds for IPSS (8 and 20) were found to be inappropriate in a treatment-seeking LUTS population, with thresholds of 15 and 27 reflecting the population better than the current cut-points of 8 and 20. The key drivers of QoL impact rank differently from the order of symptoms according to severity, meaning that decision-making should primarily consider individual item bother scores, rather than severity. When evaluating the patients at presentation, the ICIQ-MLUTS gives a much better picture of the problems that symptoms cause, particularly because it includes the key bothersome symptoms of PMD and urgency UI which need to be assessed in order to get a full picture of what influences QoL. Lack of individual item bother scores and omission of UI and PMD make the IPSS a relatively ineffective initial assessment diagnostic tool. However, the global IPSS-QoL item does provide a guide to the suitability of choosing surgical intervention for the individual patient.

Acknowledgements

The authors thank all the men who participated in the UPSTREAM trial. We also thank the participating NHS hospitals (reported elsewhere ¹⁰), in particular the Principal Investigators, Research Nurses, Clinical Trial Assistants, Administrators and all other support staff, for their dedication and hard work. The study would not have been possible without the involvement of patients and the centres, and we are extremely grateful.

We also thank the members of the independent Trial Steering Committee and Data Monitoring Committee, as well as our Trial Management Group (consisting of co-applicants, representatives from the Study Office and patient representatives) for their continued counsel and support.

This 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 Bristol Randomised Trials Collaboration (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). Study data were collected and managed using REDCap (Research Electronic Data Capture²⁵) hosted at the University of Bristol.

References

1. Irwin, D. E., Milsom, I., Kopp, Z. et al.: Prevalence, severity, and symptom bother of lower urinary tract symptoms among men in the EPIC study: impact of overactive bladder. *Eur Urol*, 56: 14, 2009
2. Kim, T. H., Han, D. H., Ryu, D. S. et al.: 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. *Int Neurourol J*, 19: 120, 2015
3. Milsom, I., Irwin, D. E.: 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*, 6: 4, 2007
4. Barry, M. J., Fowler, F. J., Jr., O'Leary, M. P. et al.: The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. *J Urol*, 148: 1549, 1992
5. Kaplan, S. A., He, W., Koltun, W. D. et al.: Solifenacin plus tamsulosin combination treatment in men with lower urinary tract symptoms and bladder outlet obstruction: a randomized controlled trial. *Eur Urol*, 63: 158, 2013
6. Simon, R. M., Howard, L. E., Moreira, D. M. et al.: Does Prostate Size Predict the Development of Incident Lower Urinary Tract Symptoms in Men with Mild to No Current Symptoms? Results from the REDUCE Trial. *Eur Urol*, 69: 885, 2016
7. Autorino, R., Zargar, H., Mariano, M. B. et al.: Perioperative Outcomes of Robotic and Laparoscopic Simple Prostatectomy: A European-American Multi-institutional Analysis. *Eur Urol*, 68: 86, 2015
8. Liao, C. H., Chung, S. D., Kuo, H. C.: Diagnostic value of International Prostate Symptom Score voiding-to-storage subscore ratio in male lower urinary tract symptoms. *Int J Clin Pract*,

- 65: 552, 2011
9. Donovan, J. L., Abrams, P., Peters, T. J. et al.: The ICS-'BPH' Study: the psychometric validity and reliability of the ICSmale questionnaire. *Br J Urol*, 77: 554, 1996
 10. Lewis, A. L., Young, G. J., Abrams, P. 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*, 5: 340, 2019
 11. Bailey, K., Abrams, P., Blair, P. S. 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*, 16: 567, 2015
 12. H., A.: Information theory and an extension of the maximum likelihood principle. In *Selected papers of Hirotugu Akaike*. NY.: Springer, 1998
 13. Youden, W. J.: Index for rating diagnostic tests. *Cancer*, 3: 32, 1950
 14. Breiman., L., Friedman., J., Stone., C. J. et al.: *Classification and Regression Trees*. New York: Routledge, p. 368, 2017
 15. Robin X., Turck N., Hainard A. et al.: Package 'pROC': Display and Analyze ROC Curves, 2020
 16. Therneau T., Atkinson B., B., R.: Package 'rpart': Recursive Partitioning and Regression Trees, 2019
 17. Reynard, J. M., Yang, Q., Donovan, J. L. et al.: The ICS-'BPH' Study: uroflowmetry, lower urinary tract symptoms and bladder outlet obstruction. *Br J Urol*, 82: 619, 1998
 18. Hofmeester, I., Kollen, B. J., Steffens, M. G. et al.: The association between nocturia and nocturnal polyuria in clinical and epidemiological studies: a systematic review and meta-analyses. *J Urol*, 191: 1028, 2014
 19. Pesonen, J. S., Cartwright, R., Mangera, A. et al.: Incidence and Remission of Nocturia: A Systematic Review and Meta-analysis. *Eur Urol*, 70: 372, 2016
 20. Chapple, C. R., Drake, M. J., Van Kerrebroeck, P. et al.: Total urgency and frequency score as a measure of urgency and frequency in overactive bladder and storage lower urinary tract symptoms. *BJU Int*, 113: 696, 2014
 21. Swithinbank, L. V., Donovan, J. L., du Heaume, J. C. et al.: Urinary symptoms and incontinence in women: relationships between occurrence, age, and perceived impact. *Br J Gen Pract*, 49: 897, 1999
 22. Agarwal, A., Eryuzlu, L. N., Cartwright, R. et al.: What is the most bothersome lower urinary tract symptom? Individual- and population-level perspectives for both men and women. *Eur Urol*, 65: 1211, 2014
 23. Jolleys, J. V., Donovan, J. L., Nanchahal, K. et al.: Urinary symptoms in the community: how bothersome are they? *Br J Urol*, 74: 551, 1994
 24. Gratzke, C., Bachmann, A., Descazeaud, A. et al.: EAU Guidelines on the Assessment of

Non-neurogenic Male Lower Urinary Tract Symptoms including Benign Prostatic Obstruction. Eur Urol, 67: 1099, 2015

25. Harris, P. A., Taylor, R., Thielke, R. et al.: Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform, 42: 377, 2009

Figures

Figure 1. Flow diagram

Figure 2. Receiver operating characteristic (ROC) curve of IPSS, ICIQ-MLUTS symptom severity and bother score plotted with patients' IPSS-QoL scores rated between "mostly dissatisfied" and "Unhappy". AUC: The areas under the curve

Figure 3. Classification tree of IPSS (A), ICIQ-MLUTS SS (B) and ICIQ-MLUTS Bo (C) indicating the thresholds between IPSS and ICIQ-MLUTS severity categories. The value of each boxes represented from the top to bottom; the outcome (0: IPSS-QoL<5, 1: IPSS-QoL≥5), the percentage of "IPSS-QoL≥5" in each categorized box and the percentage of categorized box in overall cohort.

Figure 4. Histogram of IPSS in this cohort with black and red lines indicating the current and derived bandings.

Figure 5. The distributions of scores for each separate LUTS from IPSS and ICIQ-MLUTS (symptom severity and bother) categorized by new bandings. In the ICIQ-MLUTS scores, components in the incontinence subscore not covered by IPSS are indicated by the blue rectangle.

Supplemental figure

Figure S1. Histogram of the change of IPSS-QoL score by end of study in non-surgery (n=394) and surgery patients (n=233).

Table 1. Patient characteristics in the analysis set (N=627)

		Number of patients	Mean \pm SD	Median (Min - Max)
Age		627	67.4 \pm 9.0	68 (30-91)
Uroflowmetry	Q _{max}	598	12.4 \pm 7.3	11.0 (1.4 – 58.3)
	Voided volume	604	250.3 \pm 142.3	224.0 (7.0 - 947.0)
	PVR	603	136.4 \pm 148.1	100.0 (0.0 - 1600.0)
IPSS-QoL		627	4.1 \pm 1.3	4 (0 - 6)
IPSS		627	18.8 \pm 7.1	19 (1 - 35)
ICIQ-MLUTS	Symptom severity	627	18.1 \pm 6.8	18 (3 - 40)
	Bother	627	52.3 \pm 29.2	51 (0 – 130)

SD: standard deviation; Q_{max}: maximum flow rate; PVR: post-void residual urine; IPSS-QoL: International Prostate Symptom Score Quality of Life; IPSS: International Prostate Symptom Score; ICIQ-MLUTS: International Consultation on Incontinence Questionnaires Male Lower Urinary Tract Symptoms.

ACCEPTED UNEDITED MANUSCRIPT

Table 2. Univariate regression model of IPSS-QoL scores rated between “mostly dissatisfied” and “unhappy” (N=627)

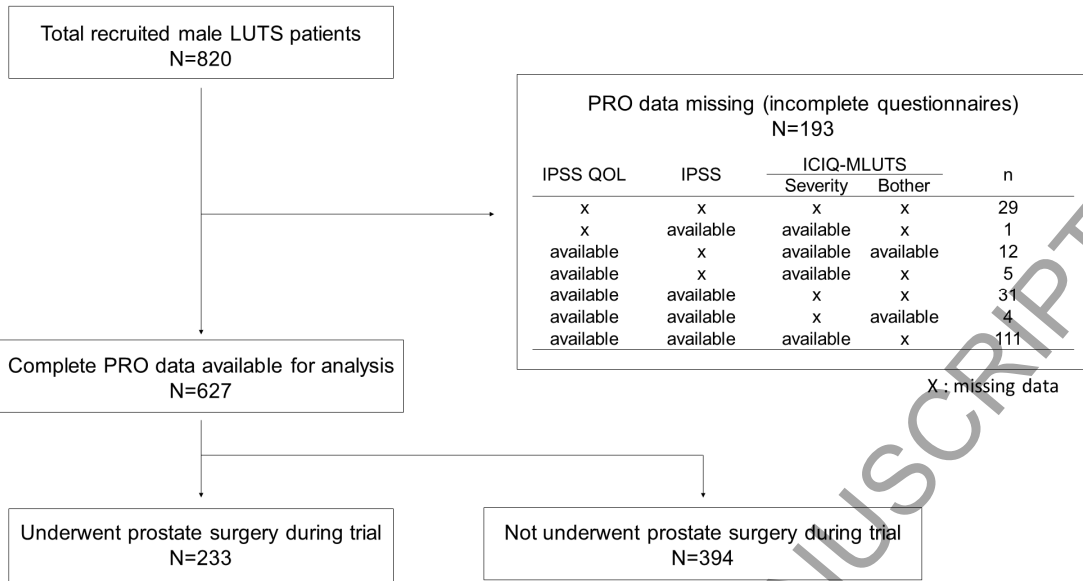
Variable	Coefficient	Lower CI	Upper CI	P value	AIC
Age	-0.009	-0.027	0.009	0.318	867.377
Q _{max}	0.003	-0.019	0.025	0.780	828.399
Voided volume	-0.001	-0.003	0	0.021	832.351
PVR	0	-0.001	0.002	0.433	835.651
Total IPSS*	0.130	0.104	0.159	<0.001	762.002
c2 (≥20 vs ≤19)	1.380	1.05	1.716	<0.001	798.494
c3 (15 - 26 vs ≤14)	1.162	0.756	1.586	<0.001	755.457
c3 (≥27 vs ≤14)	3.261	2.572	4.033		
Total ICIQ-MLUTS/severity*	0.135	0.107	0.165	<0.001	762.99
c2 (≥19 vs ≤18)	1.330	1.001	1.665	<0.001	803.266
c3 (16 - 25 vs ≤15)	0.977	0.612	1.351	<0.001	768.34
c3 (≥26 vs ≤15)	2.828	2.203	3.523		
Total ICIQ-MLUTS/bother*	0.037	0.03	0.044	<0.001	732.015
c2 (≥62 vs ≤61)	1.625	1.279	1.979	<0.001	778.136
c3 (22 - 80 vs ≤21)	1.606	1.041	2.240	<0.001	736.026
c3 (≥80 vs ≤21)	3.706	2.968	4.517		

* Items in bold indicate analysis for continuous values. c2 and c3 indicate categorized values; c2: Comparing ranges of scores in both categorized groups when distinguishing 2 groups (i.e. categorizing as mild or severe). The cut-off point predicting QoL impact (IPSS-QoL of <5 vs. ≥5) was 20 for IPSS, 19 for ICIQ-MLUTS/severity and 62 for ICIQ-MLUTS/bother. c3: Comparing ranges of scores in each categorized group, when distinguishing 3 groups (i.e. categorizing as mild, moderate or severe). Cut-off points are given in results section.

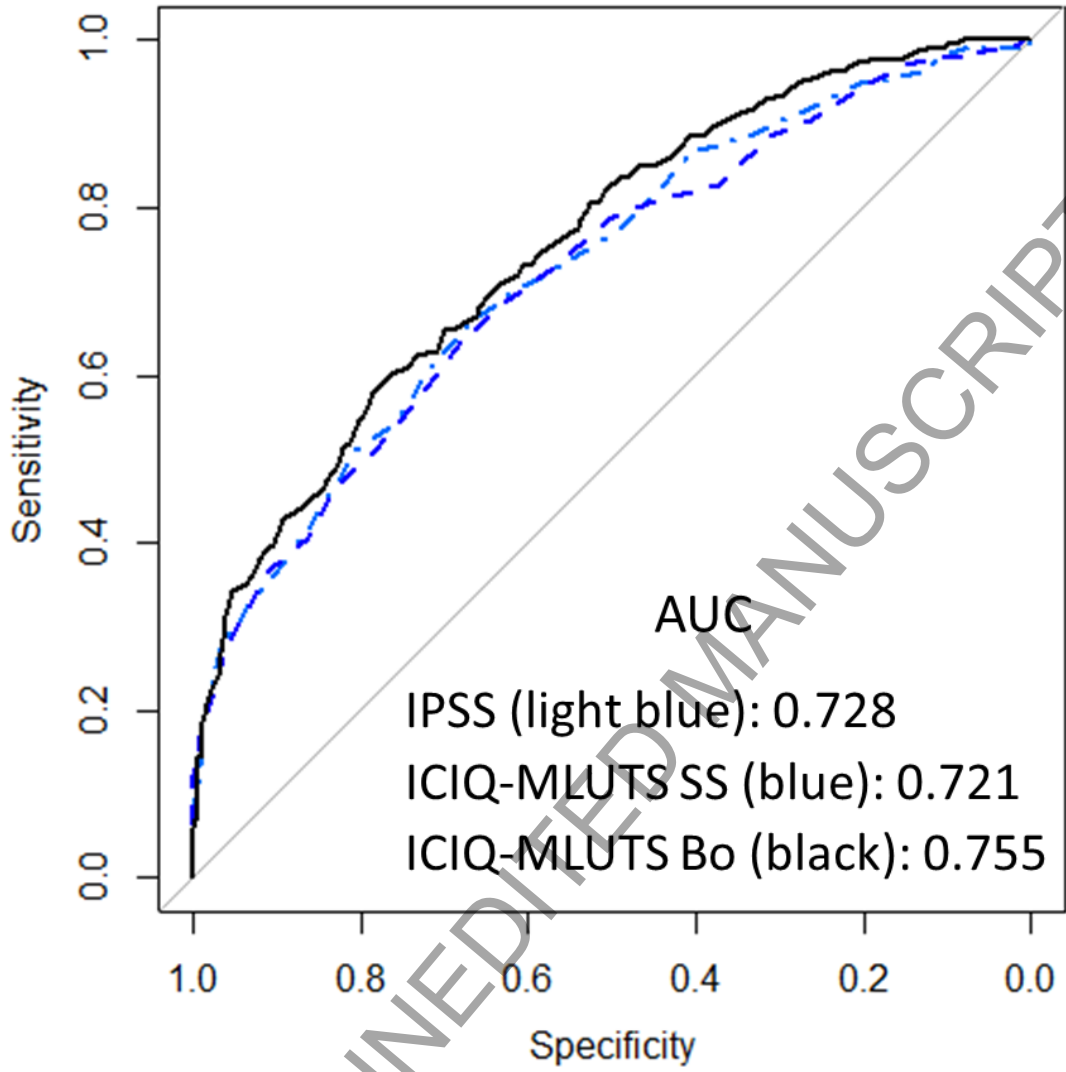
AIC: Akaike's information criteria; CI: confidence interval; ICIQ-MLUTS: International Consultation on Incontinence Questionnaires Male Lower Urinary Tract Symptoms; IPSS: International Prostate Symptom Score; PVR: post-void residual urine

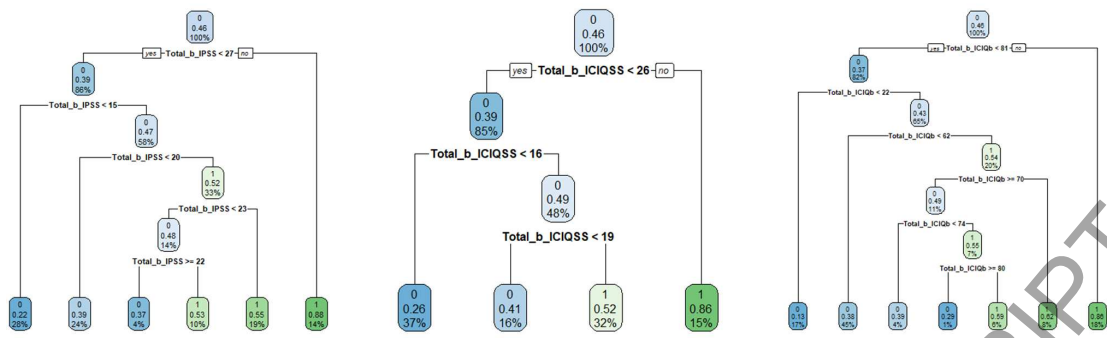
Table 3. The rankings for prevalence of severe symptoms according to symptom severity scores in IPSS and ICIQ-MLUTS and bother score in ICIQ-MLUTS

Ranking	IPSS (N=751)		ICIQ-MLUTS			
			Symptom severity scores (N=739)		Bother scores (N=643)	
	"Severe (More than half and Almost always)"	%	"Severe (Most of the time and All of the time)"	%	"Severe (8, 9 and 10)"	%
1	Weak stream	55.5	Increased daytime frequency	60.1	Nocturia	48.1
2	Nocturia	53.0	Weak stream	52.3	Urgency	45.7
3	Incomplete emptying	40.5	Incomplete emptying	41.8	Urgency UI	39.2
4	Intermittency	31.7	Urgency	39.0	Increased daytime frequency	35.5
5	Urgency	28.4	Intermittency	37.5	Incomplete emptying	35.0
6	Delay	22.9	Nocturia	26.0	Post micturition dribble	33.3
7	Increased daytime frequency	18.8	Straining	18.6	Weak stream	29.7

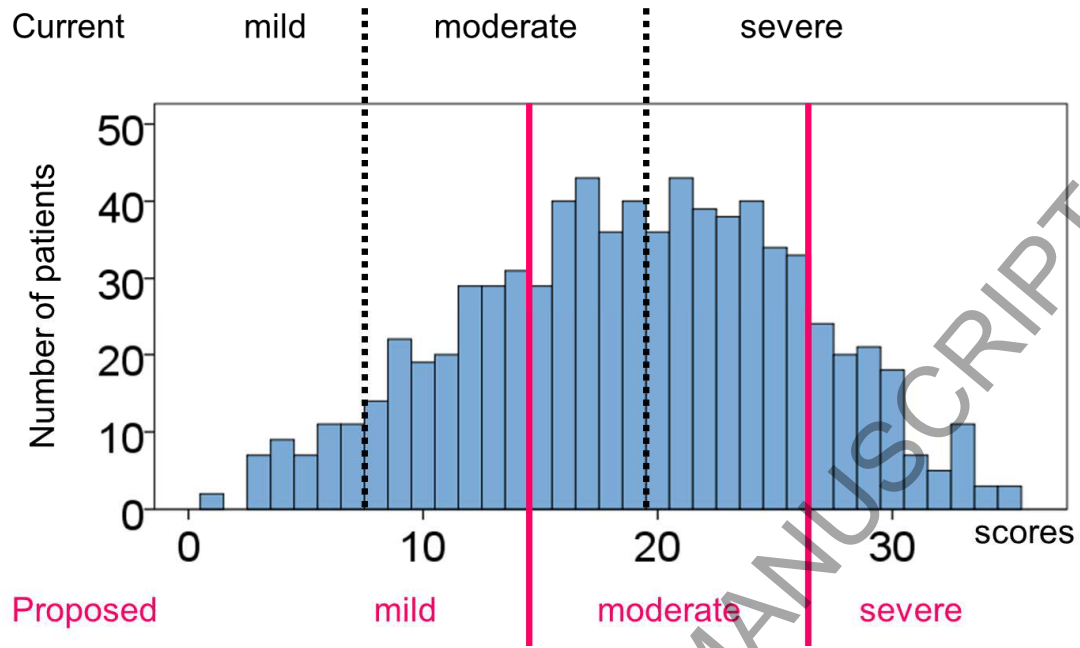


ACCEPTED UNEDITED MANUSCRIPT





ACCEPTED UNEDITED MANUSCRIPT



ACCEPTED UNEDITED MANUSCRIPT

