Urodynamics tests for the diagnosis and management of bladder outlet obstruction in men: the UPSTREAM non-inferiority RCT

Amanda L Lewis,[1,2] Grace J Young,[1,2] Lucy E Selman,[1,2] Caoimhe Rice,[1,2] Clare Clement,[1,2] Cynthia A Ochieng,[3] Paul Abrams,[4] Peter S Blair,[1,2] Christopher Chapple,[5] Cathryn MA Glazener,[6] Jeremy Horwood,[1,2] John S McGrath,[7] Sian Noble,[2] Gordon T Taylor,[8] J Athene Lane[1,2] and Marcus J Drake[4]*

[1]Bristol Randomised Trials Collaboration, Bristol Trials Centre, University of Bristol, Bristol, UK
[2]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
[3]Population Medicine, School of Medicine, Cardiff University, Cardiff, UK
[4]Bristol Urological Institute, Southmead Hospital, Bristol, UK
[5]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
[6]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[7]University of Exeter Medical School, Exeter, UK
[8]University of Plymouth, Plymouth, UK

*Corresponding author marcus.drake@bui.ac.uk

Declared competing interests of authors: Outside this work, Paul Abrams reports grants and personal fees for being a consultant and speaker for Astellas Pharma Inc. (Tokyo, Japan), and personal fees for being a consultant for Ipsen (Paris, France) and a speaker for Pfizer Inc. (New York City, NY, USA) and Sun Pharmaceutical Industries Ltd (Mumbai, India). He also reports personal fees from Pierre Fabre S.A. (Paris, France) and Coloplast Ltd (Peterborough, UK). Christopher Chapple reports being an author for Allergan plc (Dublin, Ireland) and Astellas Pharma; being an investigator for scientific studies/trials with Astellas Pharma and Ipsen; being a patent holder with Symimetics; receiving personal fees as a consultant/advisor for Astellas Pharma, Bayer Schering Pharma GmbH (Berlin, Germany), Ferring Pharmaceuticals (Saint-Prex, Switzerland), Galvani Bioelectronics (GlaxoSmithKline; Stevenage, UK), Pierre Fabre, Symimetics, TARIS Biomedical Inc. (Lexington, MA, USA), and Urovant Sciences (Irvine, CA, USA); and receiving personal fees as a meeting participant/speaker for Astellas Pharma and Pfizer. J Athene Lane was a member of the Clinical Trials Unit funded by the National Institute for Health Research during the conduct of this trial. Marcus J Drake reports being on associated advisory boards and has received grants, personal fees and non-financial support from Allergan, Astellas Pharma and Ferring Pharmaceuticals. He has also received personal fees from Pfizer.

Published September 2020
DOI: 10.3310/hta24420

Scientific summary

The UPSTREAM non-inferiority RCT

Health Technology Assessment 2020; Vol. 24: No. 42
DOI: 10.3310/hta24420

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Lower urinary tract symptoms (LUTS) are highly prevalent in men, reflecting changes in the bladder and prostate with ageing. Voiding symptoms, such as a slow stream and incomplete emptying, may indicate bladder outlet obstruction (BOO), caused by prostate enlargement. Alternatively, it may be a result of weakness of the bladder, known as detrusor underactivity (DU). Severe and bothersome LUTS are a common indication for surgery aimed at relieving BOO. The most common operation used in the UK is transurethral resection of the prostate (TURP).

The diagnostic tests used to assess men with bothersome LUTS include physical examination of the prostate, symptom score measurement, a bladder diary and flow rate testing with post-void residual scan. These give a general picture and may provide sufficient information to proceed to surgery, effectively by presuming that BOO is the underlying cause. However, urodynamics (UDS) is a test that can confirm whether BOO or DU is the cause, by measuring how much pressure is generated when passing urine. This should allow a selective approach to use of surgery, by making sure that only those men with BOO are recommended to receive an operation to relieve BOO.

Objectives

The aim of UPSTREAM (Urodynamics for Prostate Surgery Trial: Randomised Evaluation of Assessment Methods) was to determine whether a care pathway including UDS is no worse for symptom outcome than one in which it is omitted, at 18 months after randomisation. The primary clinical outcome was the International Prostate Symptom Score (IPSS) at 18 months after randomisation. The influence of UDS on surgical decision-making, as reflected in rates of bladder outlet surgery, was the key secondary outcome.

The trial addressed several other key questions:

- What is the cost-effectiveness of the two diagnostic pathways?
- What are the relative harms of UDS and the subsequent therapy?
- What subsequent NHS services are required (including repeat surgery or catheterisation for acute urinary retention) for men in each arm?
- What are the differential effects on quality of life (QoL)?

The qualitative component considered the following questions:

- What is the acceptability to and experience of participants of UDS and how satisfied are men with the diagnostic pathways?
- What are clinicians' opinions in relation to the value of UDS for male BOO?
- How does UDS affect decision-making for both surgeons and men with bothersome LUTS?
- What are the experiences and attitudes of men regarding male BOO surgery and recovery?

Methods

Design

UPSTREAM was a pragmatic, two-arm, multicentre randomised controlled trial.

Setting
This trial was set in the urology departments of 26 NHS hospitals across England ('centres').

Participants
Between October 2014 and December 2016, centres recruited men (aged ≥ 18 years) seeking further treatment, which may have included surgery, for their bothersome LUTS. Men were excluded if they were unable to pass urine without a catheter, had a relevant neurological disease, were currently undergoing treatment for prostate or bladder cancer, had previously had prostate surgery, were not medically fit for surgery and/or were unwilling to be randomised or comply with trial requirements.

Randomisation and intervention
Centre staff used a telephone- and web-based randomisation system to randomise eligible men to either a diagnostic pathway based on routine care [i.e. assessment as set out in the National Institute for Health and Care Excellence clinical guidance on male LUTS: routine care control arm [National Institute for Health and Care Excellence (NICE). *Lower Urinary Tract Symptoms in Men: Management [CG97]*. London: NICE; 2010. URL: www.nice.org.uk/guidance/cg97 (accessed 25 July 2019)] or a pathway that included UDS (i.e. routine care plus UDS: intervention arm). Centres carried out urinary flow testing and UDS in accordance with local practice; for quality purposes, equipment maintenance logs were reviewed and trace interpretation and reporting were scrutinised (≥ 10% of urinary flow and UDS traces from each centre), relative to the International Continence Society Good Urodynamic Practice requirements (Schäfer W, Abrams P, Liao L, Mattiasson A, Pesce F, Spangberg A, *et al*. Good urodynamic practices: uroflowmetry, filling cystometry, and pressure-flow studies. *Neurourol Urodyn* 2002;**21**:261–74).

A 'simple randomisation' approach was used, with no stratification or minimisation techniques. However, 'centre' was adjusted for in all analyses. Given the nature of UDS testing, and the need to access clinical data for decision-making, neither the participants nor centre staff were blinded to trial arm. The trial manager and administrative staff, although unblinded to enable individual data collection and adverse event (AE) reporting, were blinded to aggregate data. All investigators remained blinded to aggregate data throughout recruitment and analysis. The junior trial statistician had unblinded access in order to report safety and outcome data to the Data Monitoring Committee.

Primary outcome measure
The primary outcome was the patient-reported IPSS at 18 months after randomisation, using a non-inferiority design, to establish non-inferiority in symptom severity of the primary outcome, with a margin of 1 point of the IPSS scale. Data were collected via patient-completed questionnaires at baseline and at 6, 12 and 18 months after randomisation. Scores could range from 0 to 35, with higher values indicating more severe symptoms.

Key secondary outcome measure
The key secondary outcome was the number of men having surgery for their LUTS within 18 months of randomisation. Data were collected from trial case report forms completed by the centres.

Additional secondary outcome measures
Additional secondary outcome measures included the number of AEs in each arm throughout the trial, including severity, expectedness and relationship to testing and treatment. Surgery-related events (complications) were graded using the internationally acceptable Clavien–Dindo classification. All events, and classifications, were independently reviewed to ensure uniformity and check for reporting bias.

Additional patient-completed questionnaires were completed at baseline and at 6, 12 and 18 months after randomisation. These included measures of QoL (IPSS QoL), urinary symptoms severity and bother [International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms (ICIQ-MLUTS)] and sexual symptoms [International Consultation on Incontinence Questionnaire – Sexual Matters associated with Male Lower Urinary Tract Symptoms (ICIQ-MLUTSsex)]. Men who underwent UDS

also completed a satisfaction questionnaire after the procedure (International Consultation on Incontinence Questionnaire – urodynamics satisfaction) to explore patient satisfaction, details about the procedure and whether or not they would recommend it.

Maximum urinary flow rate (Q_{max}) was measured at baseline and at 18 months post randomisation (as well as at 4 months post surgery for surgical patients).

Primary analyses

The primary analyses were conducted under the intention-to-treat (ITT) principle, using multivariate linear regression. Both centre and the baseline IPSS were adjusted for in the primary analysis and results were based on the prespecified non-inferiority margin. Given the non-inferiority design, interpretation of primary analysis results focused on observed difference and 95% confidence intervals (CIs) for the between-group comparisons. When CIs lay below the non-equivalence margin, the two arms were deemed equivalent.

Secondary analyses

Secondary analyses were conducted using ITT and adjusting for centre and baseline measures (when appropriate), testing for superiority as opposed to non-inferiority, at 18 months. Several prespecified sensitivity analyses were conducted to test the robustness of the results from the statistical analyses to increase understanding of the relationship between the dependent and independent variables for the primary analysis and, in some circumstances, the key secondary analysis. Prespecified subgroups were used to test whether or not the differences between the two arms were more pronounced in certain subgroups of participants. Although underpowered, tests of interaction between the dichotomised/categorical variables and trial arm were carried out to test whether or not the treatment effect differed between subgroups. These interaction terms were added to the primary analysis model.

Cost-effectiveness analysis

Cost-effectiveness analysis was conducted from randomisation to 18-month follow-up from three perspectives: (1) NHS secondary care, (2) NHS and (3) a patient perspective. Costs were derived from resources used by men in relation to the use of secondary care, community-based NHS services and any out-of-pocket expenditure related to the treatment of their LUTS. Resource use data came from two main sources: local hospital routine data and questionnaires completed by participants at baseline and at 6, 12 and 18 months. These data were valued using UK reference costs and participant-reported costs. Quality-adjusted life-years (QALYs) were determined from the EuroQol-5 Dimensions, five-level version (EQ-5D-5L), questionnaire administered at baseline and at 6, 12 and 18 months and the EuroQol-5 Dimensions, three-level version (EQ–5D-3L), cross-walk valuation set.

Cost-effectiveness analyses under an ITT approach of the two pathways (routine care vs. routine care plus UDS), from the three perspectives in relation to QALYs, are reported. Costs and outcomes in year 2 were discounted at 3.5%. Seemingly unrelated regression was used to estimate incremental cost-effectiveness ratios and incremental net monetary benefit statistics. Uncertainty was explored using cost-effectiveness acceptability curves and sensitivity analyses.

Qualitative evaluation

In-depth semistructured interviews were conducted with different (purposive) samples to address the various objectives. This included patients who were eligible but declined to take part, participants (at various stages of their decision-making and treatment pathway) and health-care professionals.

Interviews followed topic guides developed at the start of the trial based on literature and input from the Trial Management Group. The guides were devised to ensure that the primary issues were covered across all interviews, but did not dictate data collection and were flexible to allow the introduction of new topics. Analysis was conducted in parallel with data collection, with findings from early analysis informing later data collection in an iterative process. Sampling continued until no new themes emerged. Inductive thematic analysis was used.

Results

Of the 8671 men screened for eligibility, 1482 (17%) were considered eligible, of whom 820 (55%) were randomised (427 men to the UDS arm and 393 men to the routine care arm). Sixty-seven men withdrew before the 18-month final follow-up (seven of whom requested complete data withdrawal) and 11 died (unrelated to trial procedures/treatment).

Baseline characteristics were well balanced between arms. Available outcome data were also relatively balanced between arms; however, the number adhering to the assigned diagnostic pathway differed somewhat [353/427 (83%) men randomised to UDS received UDS and 393/360 (92%) men randomised to routine care received routine care (i.e. did not undergo UDS)].

Primary and key secondary outcomes

Primary analysis results show that UDS was non-inferior to routine care for IPSS at 18 months after randomisation, with a CI within the non-inferiority margin of 1 point (adjusted difference in means –0.33, 95% CI –1.47 to 0.80). Overall, for both arms, IPSSs dropped from a mean of 18.94 ($n = 774$) to a mean of 12.86 ($n = 669$). The per-protocol analysis, along with other sensitivity analyses, gave similar results. The predicted lower surgery rate in the UDS arm was not identified, as surgery rates, overall, were much lower than expected and similar between the arms [38% (153/408) and 36% (138/384) for UDS and routine care, respectively; adjusted odds ratio 1.05 (95% CI 0.77 to 1.43)].

Clinical secondary outcomes

A total of 428 AEs were reported: 234 and 194 in the UDS and routine care arms, respectively. The number of events experienced per person was very similar for both arms, with 32% and 30% of participants in the UDS and routine care arms, respectively, experiencing at least one event. There were more cases of acute urinary retention in the routine care arm (29/289 in the routine care arm vs. 13/424 in the UDS arm, ITT). There were no apparent differences between the arms for Q_{max} (i.e. the measurement of maximum urinary flow rate) measured on or near the 18-month follow-up.

Patient-reported secondary outcomes

Urinary symptoms improved in both arms. The improvements seen were similar in both arms, although men in the UDS arm showed a greater reduction in nocturia than men in the routine care arm ($p = 0.010$). However, given the large number of secondary analyses carried out and lack of additional urinary symptom benefits, this may be a chance finding. Sexual symptoms at 18 months were very similar to baseline levels, with no differences evident between the arms.

Satisfaction with UDS was high in all men who received it, with 98% agreeing that the test was successful and 97% saying that they would have the test again.

Cost-effectiveness analysis from a secondary care perspective

The care pathway with UDS testing was more expensive than routine care by £216 (95% CI –£40 to £471). QALYs were similar between the two arms; the QALY difference was 0.006 favouring the UDS arm (95% CI –0.023 to 0.035).

Qualitative evaluation

The key findings were that UDS was acceptable to patients and valued by both patients and clinicians for its perceived additional insight into the cause, and probable best treatment of, LUTS.

Conclusions

Inclusion of UDS in the range of diagnostic tests for male LUTS results in a symptom outcome that is non-inferior to a pathway based on routine care in the UK. However, adding UDS does not affect surgical

rates for treating BOO. Including UDS in the assessment pathway leads to higher NHS secondary care costs and similar QALYs (including for the wider NHS and patient perspectives). The qualitative research identified that UDS was acceptable to patients, and it was valued by patients and clinicians for the additional insight into the cause of and treatment choice for LUTS. Overall, these results do not support the routine use of UDS for men undergoing investigation of LUTS. However, the large number of men who saw modest symptom improvements, or worsening, suggests that there may be value for the selective use of UDS, which will be explored further.

The economic analysis suggested that including UDS in the assessment of patients who present with LUTS, compared with routine care, leads to higher costs and similar QALYs from all three perspectives.

The implication for health care is the lack of justification for the routine inclusion of UDS in the diagnostic assessment of male LUTS for which surgery is being considered. Both pathways realised a substantial improvement in overall symptom scores, with no substantive difference in symptoms or surgery rates. The expressed need of patients is to seek relief of bothersome symptoms, as opposed to relief of BOO. Hence, a urology department evaluation needs to establish which LUTS are bothering the individual patient, in order to direct the focus of testing and treatment.

Recommendations for research

- The existence of a subgroup of men who suffered a deterioration in symptoms and QoL requires interpretation to establish risk factors anticipating the bad outcome with treatment and, specifically, the indications for selective utilisation of UDS in individual cases.
- The long-term outcomes for men treated for LUTS remain unclear, particularly regarding storage LUTS such as nocturia. Beyond the 18-month time point, the need for ongoing or additional treatment is probable, particularly for those men who did not get surgery and for those men undergoing surgery whose underlying mechanism was DU.
- The detailed understanding of symptom severity and bother using questionnaires warrants comparison of the main measures (IPSS and ICIQ-MLUTS).
- UDS is not the sole test used in determining whether or not a man with LUTS should be offered surgery. There are a number of possible treatments that men may be offered based on these collective tests: watchful waiting, conservative therapy, pharmacological management, minimal invasive techniques (e.g. UroLift®, NeoTract Inc., Pleasanton, CA, USA), prostate artery embolisation or surgery (e.g. TURP) and greenlight laser. The most cost-effective diagnostic strategy in this type of diagnostic pathway could be assessed using a decision model.

Trial registration

This trial is registered as ISRCTN56164274.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 42. See the NIHR Journals Library website for further project information.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 12/140/01. The contractual start date was in April 2014. The draft report began editorial review in November 2018 and was accepted for publication in February 2019. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).