NIHR | National Institute for Health and Care Research

Health Technology Assessment

Check for updates

Extended Research Article

Urodynamics tests for the diagnosis and management of male bladder outlet obstruction: long-term follow-up of the UPSTREAM non-inferiority RCT

Madeleine Clout,[1,2†] Amanda L Lewis,[1,2†] Madeleine Cochrane,[1,2] Grace J Young,[1,2] Paul Abrams,[3] Peter S Blair,[1,2] Christopher Chapple,[4] Gordon T Taylor,[5] Sian Noble,[2] Tom Steuart-Feilding,[1,2] Jodi Taylor,[1,2] J Athene Lane[1,2] and Marcus J Drake[6*]

[1]Bristol Trials Centre, Bristol Medical School, University of Bristol, Bristol, UK
[2]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
[3]Bristol Urological Institute, Southmead Hospital, Bristol, UK
[4]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
[5]Patient and public involvement representative
[6]Department of Surgery and Cancer, Imperial College London, London, UK

[*]**Corresponding author** marcus.drake@imperial.ac.uk
[†]Joint first authors

Published July 2025
DOI: 10.3310/SLPT4675

Scientific summary

Urodynamics tests for the diagnosis and management of male bladder outlet obstruction: long-term follow-up of the UPSTREAM non-inferiority RCT

Health Technology Assessment 2025; Vol. 29: No. 26
DOI: 10.3310/SLPT4675

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

Scientific summary

Background

Lower urinary tract symptoms (LUTS) are common with ageing. In men with voiding LUTS (difficulty passing urine), benign prostatic obstruction (BPO) may be treated with prostate surgery, such as transurethral resection of the prostate (TURP). However, voiding LUTS can also be caused by bladder dysfunction, for example 'detrusor underactivity' (DU). In DU, TURP may not be helpful, and could be harmful. The UPSTREAM study (Phase I) evaluated the assessment of LUTS in men who were seeking further treatment for their bothersome LUTS, including surgical intervention. Men were randomised to a care pathway which included invasive urodynamic (UDS) testing [UDS group (intervention) and routine care group (routine care)]. Following assessments, men decided on their treatment in discussion with their healthcare professional, which could have been conservative or interventional (procedures to relieve BPO). The primary outcome was patient-reported International Prostate Symptom Score (IPSS) at 18 months post randomisation; a key secondary outcome was rates of bladder outlet surgery.

The current study evaluated whether identifying the mechanism causing voiding LUTS enabled clinicians to reduce surgery rates without impairing symptom outcomes. A higher surgery rate would be anticipated in the routine care group, so identification of all men proceeding to surgery was important for the present study. However, the additional diagnostic testing potentially delayed treatment, so that the primary outcome could not reliably be captured within the 18-month time frame of the Phase I study. In the medium term, patients managed conservatively may re-present since they are likely to experience ongoing symptomatic bother, potentially resulting in a decision to proceed to additional investigation and possible surgery. Consequently, a longer-term review of these men could identify a treatment shift from the initial decision.

Objectives

The second phase of the UPSTREAM project (Phase II 'Further Follow Up') undertook a follow-up of UPSTREAM participants, aiming to determine intervention rates and outcomes at 5 years post randomisation. The objectives were to assess at 5 years post randomisation:

1. What are the symptomatic outcomes for LUTS, measured by the IPSS?
2. What are the surgery rates in the two diagnostic pathways?
3. Was additional diagnostic testing (e.g. UDS) undertaken after the completion of UPSTREAM Phase I?
4. What are the differential effects on other outcomes?
5. What is the cost-effectiveness from an NHS secondary care perspective?
6. What is the differential use of NHS resources?

Methods

The Proportionate Review Sub-Committee of the South Central – Berkshire Research Ethics Committee reviewed and approved UPSTREAM Phase II (REC reference 19/SC/0578, Integrated Research Application System ID 264738).

Participants of UPSTREAM Phase I were recontacted for one questionnaire, and routine NHS data were extracted. For the patient-reported outcome measures (PROMs questionnaire) study component, the exclusion criteria were men who:

- were not willing to be contacted about long-term follow-up; and/or
- had withdrawn study participation, or had withdrawn permission to be contacted, at the time of their 18-month follow-up time point; and/or
- did not consent and/or were not willing or able to comply with essential study procedures.

For the NHS England data extraction study component, the exclusion criteria were men who:

- were not already randomised (enrolled) to the UPSTREAM study (Phase I); or
- had withdrawn permission for the study to continue to access sections of their medical notes and NHS records, Office of National Statistics (ONS) and NHS Central registers information.

Patient-reported outcome measures (questionnaire) component; a 'Further Follow Up Patient Pack' was sent, containing the invitation letter, participant information sheet, questionnaire booklet and pre-paid return envelope, plus an online/telephone questionnaire request form. LUTS were measured with the IPSS PROM and International Consultation on Incontinence Questionnaires (ICIQs) ICIQ male LUTS and ICIQ-MLUTS-sex. EuroQol-5 Dimensions, five-level version (EQ-5D-5L) provided overall quality of life with weights used to calculate quality-adjusted life-years (QALYs). Return of the completed questionnaire indicated implicit consent. A maximum of three contact attempts was made.

Data for objectives relating to surgery rates, diagnostic testing and resource use were obtained via a one-off bespoke data extraction of Health Episode Statistics (HES) and HES-ONS linked data via NHS England:

1. HES admitted patient care (HES APC) for 5-year follow-up, spanning years 2014–22
2. HES outpatients (HES OP)
3. civil registrations of death.

Data linkage to the UPSTREAM study data set was carried out by NHS England, and the linked data were then searched for a list of 'surgery related to LUTS' and UDS codes.

All analyses were pre-specified in the statistical and health economics analysis plans. Stata version 17.1 (StataCorp LP, College Station, TX, USA) was used for all statistical analyses, and version 17.0 for all health economic analyses. The primary analysis was the IPSS total score at 5 years after randomisation. Scores were compared between the groups using multivariable linear regression, under the intention-to-treat principle, adjusting for centre and baseline IPSS. A non-inferiority margin of 1 was used and emphasis placed on the confidence intervals (CIs) around the observed difference, rather than *p*-values. The key secondary outcome was surgery rates for LUTS at 5 years, compared between the two groups using logistic regression, adjusting for centre.

Patient-reported outcome measures were also used to compare the urinary and sexual symptoms. Scores were used where applicable but individual categorical symptoms were dichotomised for ease of reporting. Ordinal scales were also analysed to ensure that dichotomisation did not mask any of the findings. Pre-specified sensitivity analyses were conducted to test the robustness of the primary analysis results and to see if any factors could have influenced the findings. Subgroup analyses were pre-specified in the Phase I study and were analysed again for Phase II.

An NHS secondary care perspective was taken in England. Care potentially related to bothersome LUTS was identified in HES APC data sets using a list of codes for treatment specialty, procedures and diagnoses. All OP consultations were included in the primary economic analysis.

Healthcare Resource Groups (HRGs), created using the latest version of the NHS Reference Costs Grouper, were used for valuation of resource data. NHS reference costs from 2020 to 2021 were used to assign a unit cost to the HRG codes at the finished consultant episode (FCE) level.

The primary economic outcome was the QALY derived from EQ-5D-5L scores which were mapped to the published UK population EuroQol-5 Dimensions, three-level version valuation set. QALYs were constructed using the area under the curve approach. For the primary economic analysis, costs and QALYs were discounted at 3.5%. Incremental adjusted mean costs and utility for 5 years were estimated using seemingly unrelated regression on multiple imputed data sets. The latter had been created using multiple imputations with chained equations and predictive mean matching. Costs and QALY estimates were adjusted for centre and baseline IPSS, and QALY estimates were adjusted for baseline utility. Estimates were combined through the net benefit framework to estimate the incremental net monetary benefit (INMB) statistic using the National Institute for Health and Care Excellence (NICE) threshold of £20,000 per QALY. Uncertainty in our INMB estimate was represented using CIs and cost-effectiveness acceptability curves (CEACs).

Results

During the Phase I study, 7 men requested complete data withdrawal and an additional 11 requested no further medical notes reviews. Hence, we collected outcomes on 802 (98%) of our original 820 cohort. NHS England failed to trace one man in the intervention group leaving 801 men in total. After a review of permissions and withdrawals, 5-year questionnaires were sent out to 550 participants. Completed questionnaires were received from 416 participants (75% overall, 211 UDS group, 204 routine care group). The proportions of men withdrawn/lost to follow-up were similar between the groups.

Using HES data, we identified 323 individuals who received UDS over their 5-year follow-up and 478 who did not. We suspect that UDS may be poorly coded in routine data, since only 265 (70%) of the 376 cases of UDS identified in Phase I study were picked up by routine NHS APC and OP data sets. There were 25 cases of UDS identified between 18-month and 5-year follow-up, mainly in the routine care group (23 vs. 2), but this may be an underestimate, given the possibility of poor coding of UD procedures.

The IPSS score (primary outcome) increased slightly from 18 months to 5 years; mean IPSS scores were 13.68 and 13.62 for the UDS and routine care groups, respectively. Non-inferiority could not be confirmed, given the wide CI. The point estimate changed direction between 18 months and 5 years, being in favour of the UDS group at 18 months and in favour of routine care at 5 years but based on very small differences in means.

Health Episode Statistics data identified 454 procedures for the surgery rates (key secondary outcome). The most common code was M653 'Endoscopic resection of prostate NEC' with 185 procedures. Furthermore, 282/291 (97%) of Phase I surgical cases, and likewise 494/499 (99%) of Phase I non-surgical cases were correctly identified, so the routine data successfully matched 776/790 (98%) of Phase I findings. There were 65 additional surgery cases with 5 years of follow-up. Also, 347 participants received surgery for LUTS; 180/415 (43%) in the UDS group and 167/386 (43%) in the routine care group. At 5-year follow-up, there had been a total of 28 deaths in the UDS group and 26 deaths in the routine care group.

For the economic evaluation, the study had high rates (97.7%, *n* = 801) of HES and survival data from the electronic health records. For the HES APC data set, 5051 records referred to a FCE that took place during the 5-year follow-up, of which 880 FCE records were relevant for the primary economic analysis. For the HES OP data set, 6102 records were used for the primary analysis. Mean NHS secondary care resource use was similar across both groups over the 5-year period. TURP surgery rates were found similar in both groups. Higher rates of UDS observed in the UDS intervention group were expected. The mean resource use for other admitted and OP procedures was low and similar across both groups. On average, participants had around one initial face-to-face OP consultation and four follow-up face-to-face consultations in a relevant treatment specialty. Both groups incurred the majority of their total secondary care costs in the first and second year post randomisation. The UDS intervention group incurred £136 and £168 more costs per person than the routine care group during the first and second years, partly explained by the extra UDS test. Costs then declined, and there were lower costs in the intervention group for the fifth year of the study (−£138). There is no evidence to indicate the COVID-19 pandemic impacted the two groups nor the 26 participating urology departments differently.

The EQ-5D-5L questionnaire was completed at all five time points for just over a third (*n* = 317, 38.66%) of all participants. Over the 5-year time horizon, adjusted incremental mean costs were slightly higher (£176.63, 95% CI −£464.06 to £817.32) in the UDS group, and adjusted incremental mean QALYs were slightly lower (−0.039, 95% CI −0.152 to 0.073). The INMB statistic was negative at −£962.62 (95% CI −£3323.54 to £1398.30), indicating the UDS group is unlikely to be the cost-effective group when applying the UK's recommended threshold of £20,000 per QALY. For all estimates, the 95% CIs were wide and crossed zero which indicated there is uncertainty in our results. A CEAC showed the INMB estimate is likely to be negative with a 21% probability of being cost-effective, at a threshold of £20,000. The sensitivity analyses did not change the interpretation.

Conclusions

The UPSTREAM Phase I study identified that a care pathway including UDS was non-inferior to routine care in terms of IPSS at 18 months, but with no reduction in the rate of surgical treatments. The extension up to a fifth year permitted the identification of all interventional treatments. This confirmed that there was no difference between the two groups in terms of the proportion receiving surgery. IPSS was captured at 5 years, identifying no difference between the groups. However, the smaller number of participants willing and able to provide an up-to-date symptom score at 5 years in the study meant it was underpowered to confirm non-inferiority. This strongly backed up the credibility of the main findings of Phase I.

The main strength of the study relates to the ability to identify the key outcome of surgery from HES data over a full 5-year time frame. The same data also represent weakness, in being unable to identify all the UD testing and reasons for OP consultations. We achieved a good response rate for questionnaires, but the length of time for an older population of men, some of whom had morbidity, meant that fewer men were available to complete PROMs at 5 years than had been at 18 months.

The study findings support progression to impact research by development of tools to optimise assessment methods, educational elements of interpreting results and quality of service delivery.

The study has considerable importance in the management of MLUTS. UPSTREAM clearly identifies that UDS did not reduce surgery rates (key secondary outcome) and that symptom outcome was non-inferior. Additionally, it is unlikely to be cost-effective. This provides a clear statement that it should not be offered routinely to all men considering surgery for LUTS. Nonetheless, the study did identify patients for whom their symptom outcome was bad (failure to improve or indeed deterioration in IPSS). Post hoc analysis showed the subgroups of men likely to benefit from surgery, and those for whom deterioration is a strong possibility. This information will be valuable for future research, offering potential benefits by improving outcomes from surgery and avoiding a proportion of the sometimes-significant associated adverse events.

Trial registration

This trial is registered as ISRCTN56164274.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 12/140/01) and is published in full in *Health Technology Assessment*; Vol. 29, No. 26. See the NIHR Funding and Awards website for further award information.

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 3.5

A list of Journals Library editors can be found on the NIHR Journals Library website

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.5 and is ranked 30th (out of 174 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

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.

Criteria for inclusion in the *Health Technology Assessment* journal

Manuscripts 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 article

The research reported in this issue of the journal was funded by the HTA programme as award number 12/140/01. The contractual start date was in April 2014. The draft manuscript began editorial review in November 2024 and was accepted for publication in December 2024. 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' manuscript 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 article.

This article presents independent research funded by the National Institute for Health and Care 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, 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, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).