NIHR | National Institute for Health and Care Research

Journals Library

Check for updates

Health Technology Assessment

Volume 29 • Issue 26 • July 2025

ISSN 2046-4924

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

Madeleine Clout, Amanda L Lewis, Madeleine Cochrane, Grace J Young, Paul Abrams, Peter S Blair, Christopher Chapple, Gordon T Taylor, Sian Noble, Tom Steuart-Feilding, Jodi Taylor, J Athene Lane and Marcus J Drake

DOI 10.3310/SLPT4675

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

This report should be referenced as follows:

Clout M, Lewis AL, Cochrane M, Young GJ, Abrams P, Blair PS, *et al.* Urodynamics tests for the diagnosis and management of male bladder outlet obstruction: long-term follow-up of the UPSTREAM non-inferiority RCT. *Health Technol Assess* 2025;**29**(26). https://doi.org/10.3310/SLPT4675

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.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

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

Abstract

Background: Lower urinary tract symptoms are common in older men and can be bothersome, leading to treatment. The UPSTREAM randomised controlled trial (Phase I) investigated whether assessment of these symptoms with invasive urodynamic testing could improve symptoms when guiding treatment options.

Objective: To assess the long-term lower urinary tract symptoms and the rates of surgery for bladder outlet obstruction in men participating in the UPSTREAM study (Phase I).

Design: Pragmatic, multicentre, parallel-group, two-group open randomised controlled study, with outcome assessors blinded to aggregate data.

Setting: Urology departments of 26 National Health Service hospitals in England.

Participants: Men ≥ 18 years, seeking further treatment for their bothersome lower urinary tract symptoms, which may include surgery, who were existing participants of the UPSTREAM study (Phase I). 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 previous prostate surgery or were unfit for surgery.

Interventions: Routine care plus invasive urodynamics (intervention) or non-invasive routine care.

Main outcome measures: The primary outcome was a patient-reported International Prostate Symptom Score 5 years post randomisation. Rates of surgery was the key secondary outcome. Patient-reported outcomes included measures of lower urinary tract symptoms, sexual function, overall quality of life and cost-effectiveness from a secondary care perspective.

Data sources: Questionnaires to participants for patient-reported outcome measures, and National Health Service England Hospital Episode Statistics and mortality data.

Results: Of 820 men randomised in UPSTREAM (Phase I) between October 2014 and December 2016, 211/427 men randomised to the intervention group completed Phase II questionnaires (49.4%) and 205/363 in the routine care group (56.5%). There was no difference found between International Prostate Symptom Scores in the two groups at 5 years (adjusted difference 0.41, 95% confidence interval −1.10 to 1.93). There was also no difference in other lower urinary tract symptoms, sexual function or quality of life. Routine data were received for 98% of men. Three hundred and forty-seven (43.3%) men with routine data available had received at least one related surgical procedure for the treatment of lower urinary tract symptoms. Over the 5-year time horizon, incremental mean costs were slightly higher (£176.63, 95% confidence interval −£464.06 to £817.32) in the intervention group and incremental mean QALYs were slightly lower (−0.039, 95% confidence interval −0.152 to 0.073) in the intervention group. This suggests that routine care is the cost-effective option.

Limitations: Around half of the men from Phase I study completed questionnaires at 5 years, although their characteristics were similar to those of non-responders, withdrawn participants or those who had died. The majority of men were of white ethnicity, so results may be less applicable to other ethnicities.

Conclusions: Five-year follow-up does not support the introduction of invasive urodynamics in reducing lower urinary tract symptoms or rates of prostate surgery, from a clinical or cost-effectiveness perspective.

Future work: This should identify if there are subgroups of patients who might benefit from the addition of urodynamics to routine care for bothersome lower urinary tract symptoms.

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.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Contents

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

List of tables

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

List of figures

List of abbreviations

A&E	accident and emergency
APC	admitted patient care
BOO	bladder outlet obstruction
BPO	benign prostatic obstruction
BTC	Bristol Trials Centre
CEAC	cost-effectiveness acceptability curve
CRF	case report form
DU	detrusor underactivity
ECDS	Emergency Care Data Set
EQ-5D-5L	EuroQol-5 Dimensions, five-level version
FCE	finished consultant episode
GP	general practitioner
HEAP	health economics analysis plan
HES	Health Episode Statistics
HRA	Health Research Authority
HRG	Healthcare Resource Group
HTA	Health Technology Assessment
ICD-10	*International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision
ICIQ	International Consultation on Incontinence Questionnaire
IMD	Index of Multiple Deprivation
INMB	incremental net monetary benefit
IPSS	International Prostate Symptom Score
ISRCTN	International Standard Randomised Controlled Trial Number
ITT	intention to treat
LTFU	lost to follow-up
LUTS	lower urinary tract symptoms
MLUTS	male lower urinary tract symptoms
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
ONS	Office of National Statistics
OP	outpatients
PIS	participant information sheet
PMG	Project Management Group
PPI	patient and public involvement
PROM	patient-reported outcome measure
PVR	post-void residual
QALY	quality-adjusted life-year
QoL	quality of life
RC	routine care
REC	Research Ethics Committee
SAP	statistical analysis plan
SUR	seemingly unrelated regression
TSC	Trial Steering Committee
TURP	transurethral resection of the prostate
UDS	urodynamics
WTP	willingness to pay

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Plain language summary

Lower urinary tract symptoms are common in older men and can need treatment. The UPSTREAM study (Phase I) aimed to help guide treatment options for these symptoms. Improvement in symptoms are measured by changes in the International Prostate Symptom Score. Men were assigned at random to have urodynamic testing or routine National Health Service care. We followed up all men for 18 months but found that some were still waiting for treatment at the end of the study.

Phase II of UPSTREAM followed up men from Phase I study until they were at 5 years from entering the study. We wanted to measure their lower urinary tract symptoms and find out how many men had undergone prostate surgery. We asked men to complete a questionnaire which included the International Prostate Symptom Score and quality-of-life measures. We also used routine National Health Service data to gather information on prostate surgery and the use of hospital services.

Of the 820 participants in Phase I, 416 completed a 5-year questionnaire and National Health Service data were available for 801. At 18 months, there was a similar reduction observed from baseline symptoms in both groups (men who did and did not have urodynamics). We saw a slight increase in symptoms between 18 months and 5 years, with no big difference between the two groups. Furthermore, 43% of men had received surgery for their urinary symptoms by 5 years. There were no big differences observed between the groups in use of hospital services or quality of life.

UPSTREAM Phase II found that there were no major differences between the two groups after 5 years. There was no difference in the number of men receiving surgery or using hospital services. There was also no difference in their urinary symptoms or quality of life.

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.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

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.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

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.

Chapter 1 Introduction

This report describes Phase II of the UPSTREAM project, the 'Further Follow Up Study'. Phase I of the project comprised a major multicentre randomised UK study, 'the UPSTREAM study' that is detailed in a separate report.[1] Full details, including study 'Background', are provided in the protocol for Phase I, which is also published.[2]

Background

In brief, lower urinary tract symptoms (LUTS) are a common feature of ageing.[3] They can be categorised as problems with voiding (passing urine) or storing urine. In men with voiding LUTS, benign prostate enlargement with ageing causes benign prostatic obstruction (BPO). For such patients, prostate surgery, such as transurethral resection of the prostate (TURP), may improve LUTS. However, voiding LUTS can also be caused by bladder dysfunction: for example, poor expulsion strength of the bladder muscle, called 'detrusor underactivity' (DU). In such men, it is hard to justify prostate surgery if BPO is not present, especially in view of potential adverse effects.

Lower urinary tract symptoms are highly prevalent in men, reflecting changes in the bladder and prostate with ageing. Voiding symptoms include a slow stream and incomplete emptying. They can indicate bladder outlet obstruction (BOO) caused by prostate enlargement, or weakness of the bladder, known as DU. Severe and bothersome LUTS are often treated with surgery aimed at relieving BOO, for example, 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 (PVR) scan. These give a general picture, but urodynamics (UDS) is the test to confirm whether BOO or DU is the cause. UDS can enable patient selection to ensure that only those men with BOO are recommended to receive an operation to relieve BOO.

The UPSTREAM study (Phase I) evaluated the assessment of LUTS in men who were seeking further treatment for their bothersome LUTS, which may have included surgical intervention.[2] Men were randomised to a care pathway which included or omitted invasive UDS testing [UDS group (intervention) and routine care (RC) group]. Following assessments, men decided their therapy in discussion with their healthcare professional, which could have been conservative (advice, medication) or interventional (procedures to relieve BPO). The primary outcome was the patient-reported International Prostate Symptom Score (IPSS) at 18 months post randomisation, and a key secondary outcome was rates of bladder outlet surgery. The IPSS is a widely used patient-reported outcome measure (PROM) measuring seven LUTS. Scores could range from 0 to 35, with higher values indicating more severe symptoms. Study recruitment spanned from October 2014 to December 2016, and the study ended in September 2018; details are published elsewhere.[2,4] The baseline characteristics and initial diagnostic testing outcomes for the study are published,[5] as is qualitative evidence regarding the attitudes to, and experience of, UDS testing from men at each end of the clinical pathway.[6] Also, 820 men (median age 68 years) were randomised. The UD group showed non-inferiority of the mean IPSSs {UDS 12.6; RC 13.1; adjusted difference at 18 months −0.33 [95% confidence interval (CI) −1.47 to +0.80]}. In the UD group, 153/408 (38%) received surgery compared with 138/384 (36%) in the RC group [adjusted odds ratio (OR) 1.05, 95% CI 0.77 to 1.43].[5] Hence, the first phase concluded that in this population, the UD randomised group was non-inferior to RC for symptom severity, but inclusion of UDS did not reduce surgical rates.

To recap, the objectives of the UPSTREAM study (Phase I) were to answer the following questions:

1. Does invasive UDS deliver similar or better symptomatic outcomes for LUTS measured by IPSS at 18 months after randomisation?
2. Does invasive UDS influence surgical decision-making, as reflected in differing surgery rates in the two diagnostic pathways?
3. What is the cost-effectiveness of the two diagnostic pathways [quality-adjusted life-year (QALY) gained at 18 months post randomisation]?
4. What are the relative harms of invasive UD tests, and surgical and conservative management?
5. What subsequent NHS services are required (including repeat surgery or catheterisation for acute urinary retention) for men in each group?

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

6. What are the differential effects on other outcomes, such as quality of life (QoL) and general health?

Implicit is the recognition that we anticipated that there would probably be different surgery rates in the two groups. Men reporting voiding LUTS undergoing UDS would have the underlying mechanism identified (BPO or DU), and those men with DU would not be recommended for surgery. In the non-UD group (RC group), there would be an assumption of BPO as the mechanism (even though in reality a proportion would actually have DU). Consequently, a higher surgery rate would be anticipated in the non-UD (RC group). This difference means that identification of all men proceeding to surgery is important, specifically for objectives #2–5, listed above.

The nature of clinical delivery in the NHS means that patient pathways for benign conditions can be protracted. For male LUTS (MLUTS), during the UPSTREAM study (Phase I), we experienced a situation whereby diagnostic tests added considerable duration to the overall pathway. Being an extra test, people in the intervention (UD) group generally had longer to wait than those in the non-UD (RC) group. In many cases, the assessment was a substantial proportion of the 18-month study duration. Hence, surgery may have taken place in the recent past at 18 months (too close for full recovery from surgery). In some cases, the assessments took longer than 18 months, so the surgery rates at 18 months may not fully reflect the decisions finally taken by doctors and patients. The UD group was at greater risk of being affected by these in both cases.

In the medium term, patients managed conservatively may re-present since they are likely to experience ongoing symptomatic bother from LUTS. Each time, the consultation may result in a decision to proceed to additional investigation and potentially undergo surgery. Consequently, a longer-term appreciation of the assessment and therapy may identify a move away from the initial decision.

The proposal for this second phase of the UPSTREAM project (Phase II 'Further Follow Up') is to undertake the follow-up of UPSTREAM participants at 5 years post randomisation, aiming to identify:

1. definitive surgery rates in the two groups
2. symptom outcomes of treatment, allowing enough time for full recovery from surgery
3. the long-term impacts of LUTS and its therapy.

The focus will continue to be on effectiveness and patient outcomes as in the original commissioning brief. Full details of the Aims and Objectives can be seen in *Aim and objectives*; in brief, we propose assessing up to 5 years post randomisation:

1. patient-reported urinary primary outcomes *(PROMs questionnaires; IPSS, ICIQ-MLUTS and ICIQ-MLUTS-sex)*
2. rates of surgery *(via NHS England)*
3. QALYs *[EuroQol-5 Dimensions, five-level version (EQ-5D-5L)]*
4. resource use for LUTS diagnostic tests and therapy *(via NHS England, where possible)*
5. cost-effectiveness from an NHS secondary care perspective *(via NHS England)*.

Rationale for 'UPSTREAM Phase II' further follow-up study

There are three principal advantages to the completion of further follow-up study:

1. Confident identification of the complete surgery rates in both study groups, overcoming the potential incomplete data for the UD group due to the greater duration needed to complete all diagnostic assessments.
2. Proper identification of the symptomatic outcomes, avoiding the detrimental effect for those men whose primary outcome of Phase I (i.e. IPSS at 18 months post randomisation) was captured during the recovery window of surgery (likely to affect the UD group more than the non-UDS, RC group).
3. Evaluation of the resource use and health economic implications in the long term; this will be a crucial differential if surgery rates are different between the two groups.

The clinical pathway in the main study (UPSTREAM Phase I) proved longer than expected such that some patients listed for surgery did not receive it by the end of their participation in the study, or underwent surgery close to the 18-month follow-up. This was an unexpected problem, representing failure to comply with NHS referral-to-treatment targets, which worsened during the study. This is potentially a fundamental issue, as the study is predicated on differential surgery rates between the two groups (we hypothesised that the UD intervention group would be non-inferior for symptoms, but at a lower surgery rate). Thus, we need scrupulous surgery data to be able to report whether surgery rates were indeed lower.

Following completion of the main study (UPSTREAM Phase I), the men with persisting voiding symptoms managed conservatively may subsequently have received surgery (with a surgical recommendation based on 'in case symptoms improve'/'nothing else to offer') in the ensuing years. This further follow-up (UPSTREAM Phase II) would address this issue, which would be a definite step to maximising impact, avoiding the 'doing the operation anyway, regardless of assessment' mentality sometimes encountered in clinical practice.

LUTS are complex and it is recognised that:

- They are potentially a long-term problem.
- They are prone to manifest a placebo response, including with surgery.
- They respond differentially to treatment (most importantly that storage LUTS, such as urgency and nocturia, are less likely to respond to surgery relieving BOO than voiding LUTS are).

Thus, this long-term further follow-up will help in understanding the sustained response, placebo impact and the behaviour of storage LUTS over time. This further follow-up will help us to state for sure whether or not UDS delivered non-inferior symptom outcomes with lower surgery rates. Furthermore, there is very little information documented about men's attitudes to the long-term experience on LUTS and its treatments, and the European Association of Urology has identified this as a priority research need.[7]

Additional insights

In addition to the above, this further follow-up will give extra information in relation to influences on a good or bad outcome. Some examples of the potential influences on outcome include:

1. Men who are bothered by a slow stream and other voiding LUTS caused by prostate enlargement should have a significant improvement with successful surgery to relieve BOO. However, only UD testing can decide if BOO is actually present; if it was not included in the diagnostic pathway, the clinician has to rely on indirect assessments to surmise whether BOO is likely to be present.
2. Nocturia is one of the main drivers for a man to present for urological assessment, and it potentially reflects a range of medical and other influences. The symptom is comparatively unlikely to improve with surgery to relieve BOO, but anecdotally some men appear to experience improved nocturia after prostate surgery.
3. Urgency is another main driver for review; some doctors think urgency may be secondary to BOO, while others consider it principally due to bladder dysfunction.
4. Post-micturition dribble is a particular nuisance symptom which affects many men to a considerable extent, despite the comparatively modest volume of urine involved.

Understanding how men can anticipate long-term change in these key symptoms according to their assessments and interventions will be a very interesting element of this 5-year follow-up.

Aim and objectives

To recap, the aim of the UPSTREAM study (Phase I) was to determine whether a care pathway including UDS (UD group) was no worse for men, in terms of symptom outcomes than one in which it was not included (RC group) at

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

18 months post randomisation.[2,4] The primary clinical outcome was measured with the widely used PROM, the IPSS, at 18 months post randomisation. A key secondary outcome was to also establish whether inclusion of invasive UDS reduced rates of bladder outlet surgery; surgical rates within 18 months of randomisation were obtained via case report forms (CRFs) completed by trained hospital staff.

It was established during the UPSTREAM study (Phase I) that there is considerable variability in the diagnostic pathway, including duration, across 26 secondary care sites across England, as well as in patient factors. We identified that several patients had not fully completed their LUTS therapy at the time of the primary outcome assessment (18 months post randomisation) or had completed it less than 6 months beforehand. Clinically, this is an insufficient time frame to offer full insight into the impacts of surgical interventions. Thus, the key aim of this further follow-up (UPSTREAM Phase II) is to establish the long-term primary outcome response, that is 5 years post randomisation, compared between groups.

Aim

To determine intervention rates and outcomes in UPSTREAM participants at 5 years post randomisation.

Objectives

To answer the following questions at 5 years post randomisation:

A. What are the symptomatic outcomes for LUTS, measured by the IPSS?
B. What are the surgery rates in the two diagnostic pathways (the relative proportion of men having surgery)?
C. Was additional diagnostic testing (e.g. UDS) undertaken after the completion of UPSTREAM Phase I? *(Where possible)*
D. What are the differential effects on other outcomes, such as symptom: severity and bother; sexual function; QoL; and general health?
E. What is the cost-effectiveness from an NHS secondary care perspective using the QALY as the economic outcome?
F. What is the differential use of NHS resources?

Chapter 2 Methods

Some of the text in this chapter is reproduced with permission from Clout *et al*.[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Study design

Further follow-up of the UPSTREAM randomised controlled parallel-group study, at 5 years post randomisation.

UPSTREAM was a pragmatic, two-group, multicentre, noninferiority study to determine the clinical and cost-effectiveness of UDS for the diagnosis and management of BOO in men with bothersome LUTS, in whom surgery was potentially being considered. The exclusion criteria were catheter use for bladder emptying, relevant neurological disease, current treatment for prostate or bladder cancer, previous prostate surgery, unfit for surgery, and/or unwilling to comply with study requirements. Men were randomised to undergo either routine evaluations in line with the applicable National Institute for Health and Care Excellence (NICE) guideline (RC group; control) or routine tests supplemented by UDS (UD group; intervention). Following the evaluation, a treatment recommendation was made by the surgeon, and the patient decided whether to accept the treatment recommendation. The primary outcome was the difference in the IPSS between the two groups at 18 months after randomisation, with a noninferiority margin of 1 point. The key secondary outcome was surgery rate (proportion of men in each study group having surgery) within 18 months of randomisation. Additional secondary outcomes included adverse events, International Consultation on Incontinence Questionnaire for MLUTS (ICIQ-MLUTS) and associated sexual matters (ICIQ-MLUTS-sex) and IPSS QoL. See UPSTREAM Phase I for additional details on the original study design.[2]

Ethics approval and research governance

The Proportionate Review Sub-Committee of the South Central – Berkshire Research Ethics Committee (REC) reviewed and approved this further follow-up study on 1 November 2019 (REC reference 19/SC/0578, Integrated Research Application System ID 264738). Health Research Authority (HRA) and Health Care Research Wales approval was issued on 4 November 2019. North Bristol NHS Trust (Sponsor) provided confirmation of capacity and capability on 13 November 2019 (NBT R&I reference: 4560). The follow-up study is included under the original study registration, International Standard Randomised Controlled Trial Number (ISRCTN) registry (ISRCTN56164274).

This study was conducted in accordance with: International Conference on Harmonisation Good Clinical Practice guidelines; UK Policy Framework for Health and Social Care Research; Data Protection Act 2018; and General Data Protection Regulation.

Informed consent was approached in a proportionate manner according to Good Clinical Practice guidelines. Details surrounding informed consent for the questionnaire and NHS England data extraction components of this further follow-up study (UPSTREAM – Phase II) are provided later in this report.

Only one non-substantial amendment (NSA01, 6 May 2022) occurred for the follow-up study after initial approvals were granted. This amendment was regarding a change to the REC/HRA listed study end date (i.e. from 30 June 2022 to 30 June 2023), to support additional time required for obtaining routine data (via NHS England) and subsequent analysis as supported by Funder, Sponsor and other relevant Committees.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Participants

Setting

For Phase I of the UPSTREAM project, urology departments of at least 26 NHS Hospitals in England were responsible for the recruitment, assessment and treatment of UPSTREAM; see details published elsewhere.[2] In Phase II, men were only contacted for questionnaires and routine data.

The central (coordinating) research team based at the Bristol Trials Centre (BTC) were responsible for coordinating and delivering the study components. North Bristol NHS Trust had oversight of data collection and responsibility for reporting through the National Institute for Health and Care Research (NIHR).

Population

Existing participants of the UPSTREAM study (UPSTREAM Phase I).

Inclusion/exclusion criteria for this further follow-up study (UPSTREAM Phase II) differ between the 'PROM' (questionnaire) component and the 'NHS England data extraction' component, due to the described differences in the consent required. These differences influence study procedures, so the inclusion and exclusion criteria for each component are outlined in the below sections for clarity.

Patient-reported outcome measures (questionnaire) study component – exclusion criteria

UPSTREAM (Phase I) participants who:

- were not willing to be contacted about long-term follow-up;
- had withdrawn study participation, or withdrawn permission to be contacted in the future for long-term follow-up, 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 of this further follow-up (UPSTREAM Phase II).

National Health Service England data extraction study component – inclusion criteria

Men randomised (enrolled) to the UPSTREAM study (Phase I).

National Health Service England data extraction study component – exclusion criteria

Patients who:

- were not already randomised (enrolled) to the UPSTREAM study (Phase I); and

UPSTREAM (Phase I) participants who:

- 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, at the time of their 18-month follow-up time point.

Study procedures

Patient-reported outcome measures (questionnaire) study component

Identification

To identify potential participants, the research team reviewed relevant data of the 820 UPSTREAM Phase I participants. A three-step process was undertaken:

1. Review responses to the UPSTREAM Phase I consent form statement *'I am willing to be contacted in the future for long-term follow-up'.*

Of those who agreed to this optional statement, the research team then checked that the participant had not:

2. withdrawn from the study, or changed relevant permissions; *nor*
3. indicated they were not willing to be contacted for long-term follow-up at the 18-month Phase I follow-up.

Men who satisfied the above points ('potential participants') were invited to take part in the questionnaire component of this follow-up study (UPSTREAM Phase II). A member of the research team completed a study-specific screening log and provided confirmation of each potential participants' outcome for this element of the further follow-up study.

Recruitment

Prior to contacting potential participants, the research team contacted general practitioner (GP) practices with a standard letter (or equivalent e-mail) stating that we intended to contact them in due course regarding the status of a patient to establish that they are alive, had capacity to take part in the study, and still lived at the registered address. Before contacting each potential participant, we attempted to contact the GP to ascertain the information. If the GP practice failed to respond, we proceeded to post the study documents as reasonable attempts to mitigate that risk had been made.

Potential participants were invited to take part by post, in the first instance. The research team sent men a 'Further Follow-Up Patient Pack' containing the Phase II study invitation letter, participant information sheet (PIS), questionnaire booklet and pre-paid return envelope, plus an online/telephone questionnaire request form. The pack did not contain a consent form. Instead, return of the completed questionnaire (paper copy, completed online or over the telephone) was taken as adequate evidence of consent ('implicit consent').

If the research team did not receive a response from a potential participant within a reasonable time, they tried calling the man and resent another pack with a reminder letter. The research team made a maximum of three contact attempts; if no response was received thereafter, then we assumed the man was lost to follow-up (LTFU).

All men who agreed to take part were logged with the research team and continued to use the unique 6-digit study (participant) identification number that was allocated to them in the main study (UPSTREAM Phase I). The research team sent a study-approved letter to the participant's GP informing them that their patient had agreed to take part in this further follow-up (UPSTREAM Phase II).

Data collection

Return of a completed questionnaire booklet (or at least provision of critical data, such as the IPSS and EQ-5D-5L) marked the end of the participant's direct involvement. If necessary, we offered a shortened version of the questionnaire booklet to ensure that these critical data were obtained. Participants could also complete the questionnaires online or via telephone. Upon receipt of a completed questionnaire booklet, the research team offered the participant a £20 gift voucher. Men were also sent participant newsletters telling them about the study, including progress and results once available.

COVID-19 impact

Questionnaire completion rates were affected by the onset of the COVID-19 pandemic. Besides restrictive access to university premises for staff, there were delays in postal returns during COVID-19, as some men were shielding and/or facing other health priorities. We continued to use a range of contact methods, including telephone and/or online completion of questionnaires where feasible. We monitored closely and adapted our contact methods and timings (for initial invitations and reminders) accordingly, in line with our flexible protocol.

National Health Service England – data extraction study component

Identification and consent

The research team reviewed relevant data of all 820 UPSTREAM participants; these data were extracted from the existing UPSTREAM databases, and from manual checking of the UPSTREAM Phase I consent forms and withdrawal/change of permissions forms, if required.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Men who enrolled in the UPSTREAM study (Phase I) provided written informed consent, which included the statement.

> *I agree that information relevant to the UPSTREAM study may be collected from my hospital and NHS records, including Office of National Statistics (ONS), NHS central registers and the Health and Social Care Information Centre, and that this data may be used to follow my ongoing health after the study.*

The team checked that the participants had not withdrawn permission for the study to continue to access sections of their medical notes and NHS records, ONS and NHS Central registers information, at the time of their 18-month time point; data extraction via NHS England was only requested for men meeting these criteria.

A member of the research team completed the study-specific screening log and provided confirmation of the man's outcome for this element of the further follow-up study.

Data collection

Additional clinical and resource use data were collected via a one-off bespoke data extraction of Health Episode Statistic (HES) and HES-ONS data via NHS England. There were three elements to the NHS England data extraction request:

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

Data linkage was carried out by NHS England and involved linking the HES and HES-ONS mortality data sets with the UPSTREAM study data set. The research team at the University of Bristol provided NHS England with identifiable data for all patients involved in the UPSTREAM study who have provided consent for their health records to be accessed. These identifiers included: study identification number, NHS number, forename, surname, date of birth, gender and current postcode as held on study records, plus dates of initial randomisation and 5 years post randomisation.

The data extraction was originally planned to take place in 2022. Although the application for data was submitted in September 2021, there were significant delays to the receipt of data. The application was approved in January 2023, the Data-Sharing Agreement fully signed in March 2023 and data received in May 2023. The delay in receipt of these key data resulted in the study end date being extended to allow for full reporting of UPSTREAM Phase II. Data were downloaded onto a secure encapsulated virtual machine secure at the host institution (University of Bristol). Only the study statistician (GY) and study health economist (MCo) had access to this machine.

Data manipulation

There were various steps to obtaining the relevant information from the received NHS England data (*Figure 1*). Firstly, both the OP and APC data sets were used and operations were filtered by treatment specialty (TRETSPEF). There were 10 codes used to filter operations, and these were derived in Phase I.[1] Duplicate episodes and empty operations were then removed from the data set.

The Chief Investigator (MD) went through the list of operation codes, without any identifiers or randomised allocations, to generate a list of 'surgery related to LUTS' codes. Participants could have zero or multiple entries in the APC and OP data sets and each entry/spell contained between 1 and 24 operations. All of these were searched for 23 codes which the Chief Investigator had deemed 'surgery related to LUTS': M494, M495, M611, M613, M618, M642, M651, M653, M654, M655, M662, M668, M672, M678, M681, M682, M683, M705, M707, M708, Y089, Y111 and Y171. An additional eight codes were held in reserve, which were deemed 'possible surgery related to LUTS': Y531, Y534, Y535, Z421, Z422, Z424, Z425 and Z429. These were checked by the statistician to see if they indicated any other individuals who had had surgery.

Urodynamics was identified in a similar way, by searching for three codes that explicitly mentioned 'urodynamics' in the title: M474, M482 and U264.

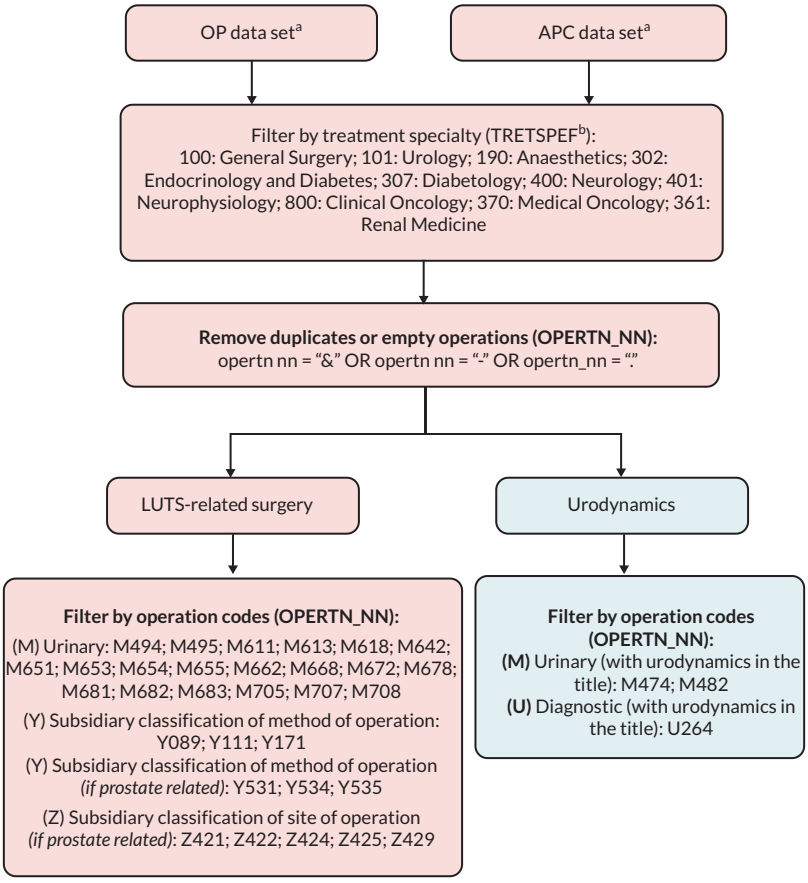

FIGURE 1 Finding UDS cases and identifying surgical procedures, for LUTS, in routine data sets. a, Data provided by NHS England (formerly NHS Digital). b, As used in Phase I: Lewis *et al.*[1] Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The figure includes minor additions and formatting changes to the original text.

Planned interventions and allocation (randomisation) to study groups

UPSTREAM Phase II refers to the further follow-up of existing UPSTREAM Phase I participants, specifically at 5 years post randomisation. As such, no new or subsequent interventions, nor allocation (randomisation) to study groups, were required. Details of the study (UPSTREAM Phase I) interventions and allocation to study groups are available in the separate protocol (Protocol Phase I, version 4, 29 September 2016), and Health Technology Assessment (HTA) Report which are also published.[1,2]

Outcome measures

Outcomes

- LUTS 5 years post randomisation, measured with the IPSS PROM (objective A). The IPSS is validated,[9] well known and widely used in the NHS.
- Alongside the IPSS (including the QoL measure), measures from selected ICIQs were used, collectively giving a sensitive and comprehensive assessment of QoL, LUTS severity/bother, sexual function and general health (objective D). The following measures were used:
  - IPSS – including QoL;
  - ICIQ-MLUTS^; and
  - ICIQ sexual function in MLUTS (ICIQ-MLUTS-sex^).

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

^*Copies of ICIQ materials can be requested from* http://iciq.net.

- The EQ-5D-5L was used to provide the QoL weights used to calculate QALYs (objective E).
- Data for objectives B (surgery rates), C (diagnostic testing) and F (resource use) were obtained via a one-off bespoke data extraction of HES and HES-ONS linked data via NHS England.

The components and timing of outcome measures are shown in *Table 1*.

TABLE 1 Measurement outcomes table: components/timing

	Five years post randomisation[a]	2023[b]
PROMs		
IPSS (including QoL)	●	N/A
EQ-5D-5L	●	N/A
ICIQ-MLUTS	●	N/A
ICIQ-MLUTS-sex	●	N/A
NHS England data extraction		
NHS England (HES APC, HES OP, Civil Registration – Deaths)	N/A	◊[c]

N/A, not applicable.
a Each UPSTREAM participant was followed up as close as possible to their 5 years post-randomisation time point. Given the long-term nature of this further follow-up, however, data from 4 years to 6 years post randomisation were accepted. If applicable, available data outside this range were reviewed and considered by the Study Statisticians and Trial Steering Committee (TSC).
b A one-off data extraction of HES and HES-ONS linked data (NHS England) took place in 2023 to include the last participant's 5-year post-randomisation time point.
c NHS England bespoke data extraction undertaken by the research team.

Non-inferiority margin

For Phase I, a non-inferiority design was chosen to best assess whether or not men randomised to receive UDS would have PROMs (urinary symptoms) that were better than RC or no worse than an acceptable number of symptoms. A lower IPSS in the intervention group, or IPSS of no more than one point higher, was hypothesised as suggesting non-inferiority. The rationale for this is described in the Phase I HTA report. The same margin is used for both Phase I and Phase II primary outcomes.

Sample size

Calculation of the Phase I sample size was based on the consideration that men randomised to UDS should have symptoms that are non-inferior to those who are randomised to RC. A full description can be found in the Phase I HTA report. The Phase II sample size is a subset of the main study population, as described.

Blinding

Given the nature of UDS testing and knowledge of results underpinning treatment decisions, group allocation was not concealed from men (nor their clinical care team) during the UPSTREAM study (Phase I). The study manager and administrative staff, although unblinded to enable individual data collection, were blinded to aggregate data.

Two statisticians continued to support this study during the second phase and the senior statistician (and co-applicant) remained blinded throughout. The secondary statistician had unblinded access to the data to report outcome data as required. The health economist was blinded when cleaning data, but unblinded when conducting the analysis. Other

members of the study team remained blinded to aggregate data. The protocol, statistical analysis plan (SAP) and health economics analysis plan (HEAP) were approved by the Project Management Group (PMG) and Trial Steering Committee (TSC) prior to the start of any analyses.

Statistical methods

All analyses were based on those pre-specified by the SAP, which was finalised in March 2023, before the analysis of the 5-year follow-up data. The SAP follows the original plan used in Phase I,[4] with a few modifications due to the routine data element. Stata version 17.1 (StataCorp LP, College Station, TX, USA) was used for all statistical analyses. Consolidated Standards of Reporting Trials reporting guidelines were followed.[10]

Primary analysis

As with UPSTREAM Phase I, the primary analysis was the IPSS total score, collected at 5-year follow-up after randomisation, using patient-reported questionnaires. Scores were compared between the groups using multivariable linear regression, under the intention-to-treat (ITT) principle, adjusting for centre and baseline IPSS. A non-inferiority margin of 1 was used and emphasis placed on the CIs around the observed difference, rather than *p*-values. The null and alternative hypotheses follow on from the Phase I study:

H_0: RC leads to IPSSs that are at least 1 point lower, on average, than in the UD group (higher IPSSs signify a poorer outcome).
H_1: UDS is non-inferior to RC, with IPSSs that are lower than in the RC group or < 1 point higher.

Secondary analyses

The key secondary outcome was surgery rates, for LUTS, at a 5-year follow-up. The source of these data in Phase I was case note review forms, whereas in this Phase II study, the data were collected using routine data from OP and admitted patient HES data sets. The proportion of men receiving surgery was estimated using the filtering methods illustrated in *Figure 1* and 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 men were expecting at 5-year follow-up: IPSS QoL, ICIQ-MLUTS and ICIQ-MLUTS-sex. Scores were used where applicable but individual categorical symptoms were dichotomised for ease of reporting (as categorised in Phase I)[1]:

- daytime frequency [voiding < 9 (0) vs. 9 or more times per day (1)]
- nocturia [voiding < 2 (0) vs. 2 or more times per night (1)]
- erection quality [normal rigidity (0) vs. reduced/severely reduced/no erection (1)]
- ejaculation quality [normal quantity (0) vs. reduced/significantly reduced/no ejaculation (1)]
- painful ejaculation [no (0) vs. slight/moderate/severe pain or discomfort (1)]
- urinary symptoms affected sex life [not at all (0) vs. a little/somewhat/a lot (1)].

Ordinal scales were also analysed to ensure that dichotomisation did not mask any of the findings.

Sensitivity analyses

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.

- Per-protocol analysis: The primary analysis was repeated, analysing men who had been treated according to the treatment group they were allocated, that is those receiving UDS after allocation to RC and those receiving RC after allocation to UDS were excluded from the analyses. This analysis was repeated, once to utilise the allocation variable collected in Phase I and, secondly, to utilise the allocation variable collected in routine data. Although only one per-protocol analysis was planned, given the discrepancy in UDS figures between Phase I CRF data and Phase II routine data, it seemed sensible to analyse both study data.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

- As treated analysis: The primary analysis was repeated, analysing men according to the treatment they received, regardless of allocation.
- Repeated measures: A repeated-measures approach was used to factor in the short- and long-term effects of each pathway. It includes data collected at baseline, 6, 12 and 18 months as well as 5 years.
- Imputing missing data: Following the methods used in Phase I, missing IPSS data, assumed to be missing at random, were imputed under conservative assumptions. Multiple imputation by chained equations was carried out, utilising variables that were related to the 18-month IPSS, treatment-specific, prespecified as important confounders or were predictive of its missingness. These variables included group of the study (binary); whether or not UDS was received according to Phase I data and Phase II data (binary); whether or not men received surgery according to Phase I and Phase II data (binary); age (continuous); presence of comorbidities (binary); centre (categorical); Q_{max} at 18 months (continuous); and all IPSS individual items at each time point (continuous). All continuous variables that were imputed used predictive mean matching, drawing from the five closest observations. Age and centre were both complete and were, therefore, not imputed. The prespecified seed of 648 (from Phase I) was used for the imputation.[1]
- Adjustment for pre-specified confounders: The following variables were pre-specified as 'clinically important' and will be adjusted for in a sensitivity analysis: centre, age, comorbidities and symptom severity (at baseline).
- Time from surgery: Using the data obtained from routine sources, the time between surgery and the 5-year follow-up time point was computed and adjusted for.

Subgroup analyses

Subgroup analyses were pre-specified in Phase I and were analysed again for Phase II. Formal tests of interaction between the dichotomised variables and treatment pathway were carried out to test whether treatment effect differs between patients. These subgroup analyses were applied to the primary analysis (IPSS score) and the main secondary analysis (surgery rates):

- age (continuous)
- flow rate (> 12 ml/second vs. ≤ 12 ml/second)
- maximum voided volume (< 200 ml vs. ≥ 200 ml)
- nocturia (< 2 times per night vs. ≥ 2 times per night)
- severity of storage LUTS using the median found in Phase I (less substantial ≤ 9 vs. more substantial > 9).

Chapter 3 Results

Recruitment

Participant flow in Phases I and II

As reported previously,[1] of the 8671 men screened for eligibility in the Phase I study, 820 were randomised to the UPSTREAM study: 427 in the UDS group and 393 in the RC group. Randomisation took place between October 2014 and December 2016, with 5-year follow-ups commencing between October 2019 and December 2021.

Five-year follow-up of clinical outcomes

During the 18-month follow-up period of the Phase I study, 7 men requested complete data withdrawal and an additional 11 requested no further medical notes reviews (*Figure 2*). Therefore, this allowed us to collect data on clinical outcomes (e.g. surgery rates, death) on 802 (98%) of our original 820 cohort. NHS England failed to trace 1 man in the intervention group leaving 801 men in total.

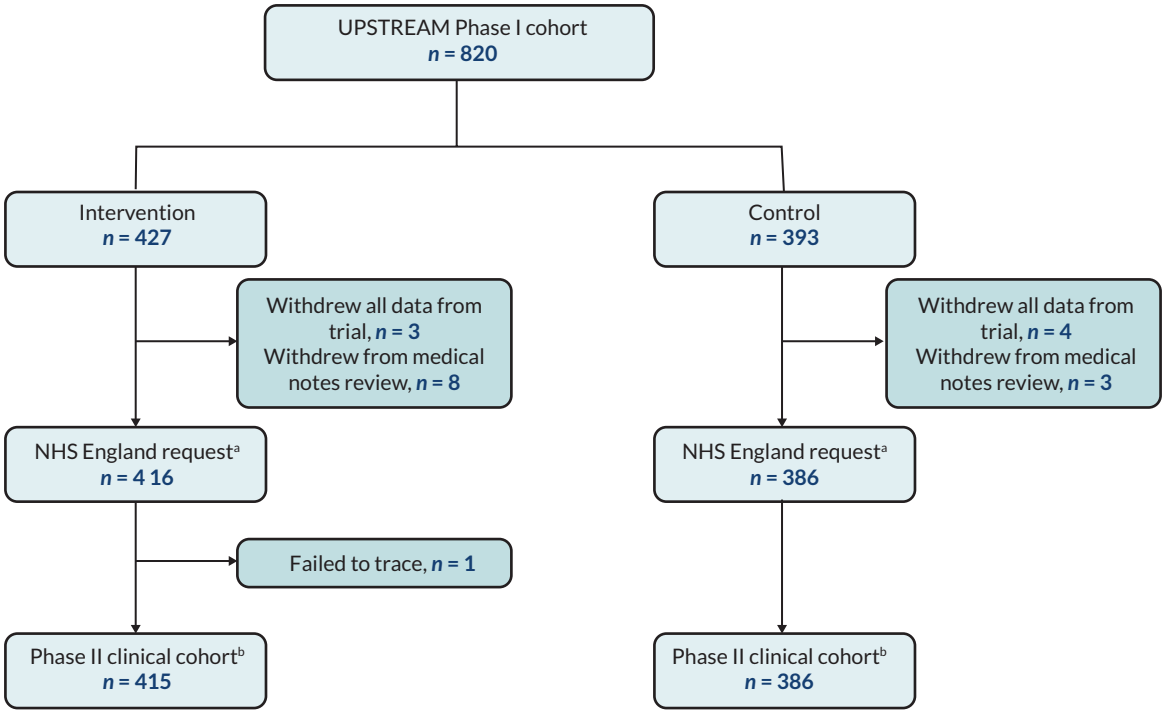

FIGURE 2 Participants eligible for the medical notes' extraction. a, Data were requested, from NHS England, for all participants who had allowed us to access their health records at 5-year follow-up. b, NHS England data extraction: The data extraction included individual level data about participants' relevant: inpatient stays; OP attendances: including procedures; radiology episodes; and cause of death (where applicable). Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The figure includes minor additions and formatting changes to the original text.

Five-year questionnaire follow-up

During the 18-month follow-up period in Phase I, there were 39 and 28 withdrawals by men, for the UD and RC groups, respectively. There were more deaths in the UD group (n = 9) than in the RC group (n = 2). There were additional withdrawals on, or shortly after, the end of the follow-up, leading to 103 withdrawals from the Phase I study.

Participants had two opportunities to decline follow-up in a future study. At Phase I baseline, at the point of consent, participants could state 'I am willing to be contacted in the future for long-term follow up' and 44 participants left this blank which was taken as a refusal. At Phase I 18-month follow-up visits, research staff were asked to confirm in

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

the CRF that 'The patient is willing to be contacted in the future for longer-term follow-up', where an additional 105 participants were identified as not eligible for future contact.

As summarised in *Figure 3*, we identified 570 men who were 'contactable' about Phase II of the project. Twenty participants were then not approached, as the site teams were aware of the participants' death or ill health. Five-year questionnaires were sent out to 550 participants.

Twenty-one participants declined to complete the questionnaire (7 in the UD group and 14 in the RC group) and 5 participants per group were identified as having died via returned post. Completed questionnaires were received from 416 participants, while 90 did not return the questionnaire after 3 contact attempts (deemed LTFU). Death notifications were based on correspondence with either the participant's GP or family member. A more robust measure of deaths, using NHS England ONS data, can be found in *Table 2*.

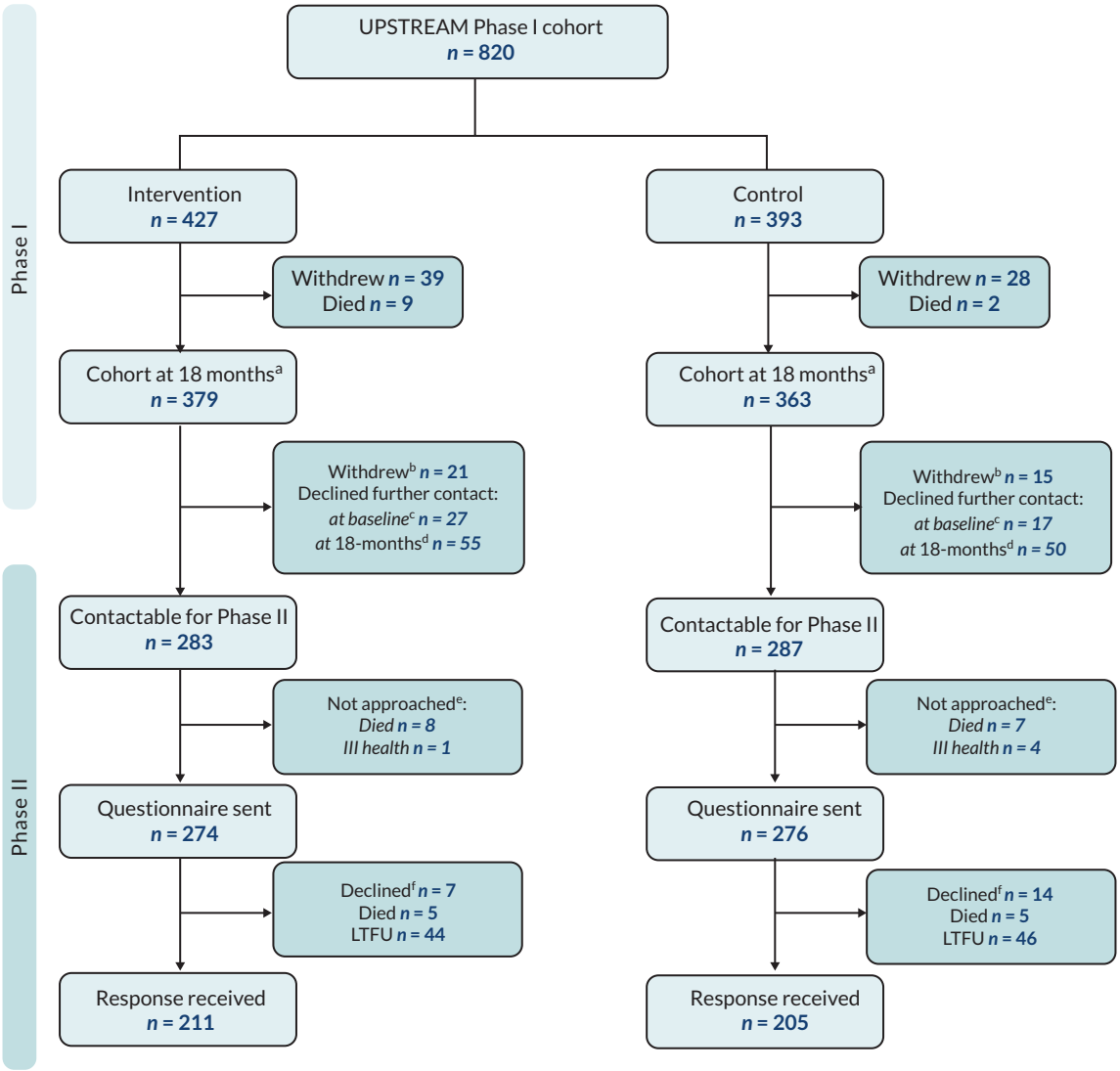

FIGURE 3 Initial screening of participants eligible for Phase II questionnaires. a, These may not match figures quoted in the Phase I publications as participants who withdrew may have allowed us to collect surgical information, via notes review. b, Chose to withdraw at the end of their 18-month follow up, or shortly after. c, At the point of consent, participants were asked to (optionally) tick whether they agreed with the following statement: 'I am willing to be contacted in the future for long-term follow up'. d, In the 18-month CRF, research site staff completed the following question, in consultation with the participant: 'The patient is willing to be contacted in the future for long-term follow up'. e, Reasons for not approaching a participant were only known if the participant's GP, or family member, advised the study team. f, Reasons were reasonably balanced between arms and the most common reasons were: 'ill health' and 'no longer interested'. N.B. The 'control' group is referred to as the 'routine care' group throughout this report. Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The figure includes minor additions and formatting changes to the original text.

TABLE 2 Withdrawals, deaths and loss to follow-up, in UPSTREAM Phases I and II

Variable	UDS *n*/427 (%)	RC *n*/393 (%)	*p*-value[a]
Withdrawals			
Withdrew within Phase I	39 (9%)	28 (7%)	0.294
Death within Phase I	9 (2%)	2 (1%)	0.047
Declined further contact at Phase I baseline/18 months	82 (19%)	67 (17%)	0.424
Withdrew within Phase II	28 (7%)	29 (7%)	0.644
Death within Phase II	13 (3%)	12 (3%)	0.994
LTFU (Phase II Questionnaires)	44 (10%)	46 (12%)	0.522
Not requested/traced by routine data	12 (3%)	7 (2%)	0.328

a Chi squared test

As shown in *Table 2*, the proportions of men withdrawn/LTFU were similar between the groups. There were slightly more deaths in the UD group, in Phase I, as presented previously.[1] None of these deaths were considered to be related to the study procedures.

Baseline characteristics

Baseline characteristics were compared between those who were included in the 5-year analysis for both randomised groups versus those not represented in the 5-year analysis (see *Appendix 1*, *Table 23*). On inspection, one individual returned a blank questionnaire; they were grouped into the 'Not included' column of this table.

When comparing the characteristics of those men analysed at 5 years with those who withdrew, died or were LTFU (not analysed), we see very similar characteristics (*Table 3*). There were more men with comorbidities among those not analysed (71%), which was not surprising given that they included (1) those who withdrew (where ill health was one of the main reasons for withdrawal) and (2) those who died.

TABLE 3 Baseline characteristics, by group, for those who completed the 5-year questionnaire

	n[a]	UDS	*n*[a]	RC
Age				
Age at randomisation (years)	211	67.14 (8.23)	204	67.53 (7.25)
Ethnicity				
White	206	188 (91%)	198	184 (93%)
Other ethnic group	206	18 (9%)	198	14 (7%)
IMD score (based on postcodes)				
Median IMD score 2015 (IQR)[b]	205	14 (8–21)	199	14 (7–22)
Quintile 1 *(most deprived)*		20 (10%)		23 (12%)
Quintile 2		35 (17%)		29 (15%)
Quintile 3		51 (25%)		48 (24%)
Quintile 4		52 (25%)		42 (21%)
Quintile 5 *(least deprived)*		47 (23%)		57 (29%)

continued

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 3 Baseline characteristics, by group, for those who completed the 5-year questionnaire (*continued*)

	n[a]	UDS	*n*[a]	RC
Clinical baseline characteristics				
Comorbidities	210	127 (60%)	199	136 (68%)
Additional tests				
PSA test	210	26 (12%)	203	30 (15%)
Cystoscopy	210	18 (9%)	203	9 (4%)
Urinalysis	210	21 (10%)	203	28 (14%)
Kidney ultrasound	210	4 (2%)	203	6 (3%)
Cytology	210	2 (1%)	203	1 (< 1%)
Prostate volume measurement	210	9 (4%)	203	4 (2%)
Urea and electrolytes	210	7 (3%)	203	6 (3%)
IPSS: symptom severity at baseline				
Incomplete emptying	210	2.50 (1.67)	203	2.84 (1.71)
Frequency	210	3.30 (1.26)	203	3.51 (1.28)
Intermittency	210	2.57 (1.62)	203	2.53 (1.59)
Urgency	208	2.53 (1.59)	203	2.58 (1.65)
Weak stream	210	3.14 (1.57)	202	3.28 (1.56)
Straining	209	1.47 (1.47)	203	1.73 (1.58)
Nocturia	210	2.43 (1.23)	202	2.73 (1.28)
Total IPSS	207	17.97 (6.67)	201	19.14 (7.28)
IPSS QoL	210	3.93 (1.32)	202	4.05 (1.25)
ICIQ MLUTS				
Voiding score	205	8.86 (3.98)	204	9.57 (4.22)
Incontinence score	202	4.70 (3.09)	198	5.15 (3.07)
Daytime frequency (> 8 times)	205	78 (38%)	200	81 (41%)
Nocturia (> 1 times per night)	204	152 (75%)	200	160 (80%)
ICIQ MLUTS – sexual matters				
Erections (reduced or none)	203	144 (71%)	198	146 (74%)
Ejaculation (reduced or none)	199	157 (79%)	195	159 (82%)
Painful ejaculation (yes)	190	31 (16%)	188	41 (22%)
Urinary symptoms affected sex life	197	138 (70%)	195	125 (64%)

IMD, Index of Multiple Deprivation; IQR, interquartile range; PSA, prostate-specific antigen.
a The number with available baseline data.

Note
Data are mean (SD) or *n* (%), unless otherwise stated.
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/.

Out of those who completed the questionnaire, 211 were originally allocated to the UD group and 204 were in the RC group. The UD group men had slightly fewer symptoms (see *Table 3*), as IPSS and ICIQ symptom scores were generally higher in the RC group and there was a greater proportion of patients with comorbidities in the group. However, there were no differences between the two groups, which met the criteria of 10% or 0.5 standard deviations (SDs), so the pre-specified baseline-adjusted sensitivity analysis was not carried out.

Adherence: receipt of urodynamics

Using the routine data, we identified 323 individuals who had received UDS over their 5-year follow-up and 478 who did not receive UDS (*Table 4*).

The agreement between Phase I and Phase II cases suggests that UDS may be poorly coded in routine data as, of the 376 cases of UDS identified in Phase I, only 265 (70%) were picked up by routine NHS APC and OP data sets. Of the 425 men identified as not receiving UDS in Phase I, 367 (86%) were matched as not receiving UDS in Phase II. Of the 58 men that were found to have received it, 37 received it within 18 months of randomisation, contradicting the findings found in Phase I. *Table 5* looks at the number of men receiving UDS at 18 months and 5 years, based on either the CRF data obtained in Phase I or the routine data collected in Phase II, excluding those not included or traced in the Phase II routine data set.

TABLE 4 How well do NHS routine data match the Phase I CRF data, for UD procedures?

	Phase II (NHS data) UDS status for 0–5 years			
Phase I (CRF) UDS status (0–18 months)	**No UDS (0–18 months) *n* (%)**	**UDS conducted (0–18 months) *n* (%)**	**Additional UDS (18 months to 5 years) *n* (%)**	**Total**
UDS conducted	111 (23.2)	261 (match) (87.5)	4 (match)[a] (16)	376
No UDS	367 (match) (78.6)	37 (new) (12.4)	21 (new) (84)	425
Total	478	298	25	801

a Includes four UDS cases which were outside the 548-day window but also captured in Phase I CRFs (late follow-ups).

Note
Light blue shading = Phase II NHS data matching with Phase I CRF data.
Dark blue shading = Phase II NHS data which is new from Phase I CRF data.
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/.

TABLE 5 Proportion receiving UDS, difference between the groups

Variable	UDS *n*/415 (%)	RC *n*/386 (%)	OR[a] (95% CI)	*p*-value[a]
UDS outcome (18 months) – using CRFs in Phase I				
UDS	348 (84%)	28 (7%)	66.41 (41.71 to 105.73)	< 0.001
No UDS	67 (16%)	358 (93%)		
UDS outcome (18 months) – using routine data on OP and admitted patient data sets				
UDS	249 (60%)	49 (13%)	10.32 (7.21 to 14.76)	< 0.001
No UDS	166 (40%)	337 (87%)		
UDS outcome (5 years) – using routine data on OP and admitted patient data sets				
UDS	251 (60%)	72 (19%)	6.67 (4.83 to 9.22)	< 0.001
No UDS	164 (40%)	314 (81%)		

a Not adjusted as centre was a perfect predictor in some analyses.

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Using the CRF data, those in the UDS group were at a 66 times greater odds of receiving UDS than those in the RC group, at 18 months follow-up. This reduced to 10 times greater odds over 18 months when using the routine data. Focusing on the routine data, there were 25 cases of UDS between 18-month and 5-year follow-up. The majority of new UD procedures (picked up between 18-month and 5-year follow-up) were in the RC group (23 vs. 2). This reduced the OR even further to six times great odds of receiving UDS over 5 years in the UD group, compared with RC.

Primary outcome

As reported previously,[1] at 18-month follow-up, it was concluded that the UD group was non-inferior to the RC group with respect to the IPSS score; difference in means −0.33 (95% CI −1.47 to 0.80). The non-inferiority margin was pre-specified as 1 point. As the upper CI was below 1, non-inferiority was concluded.

Over the first 18 months of follow-up (Phase I), the IPSS score dropped over time as treatment schedules were introduced. *Figure 4* shows what the IPSS total score trajectories looked like over time, for men who completed a questionnaire at all five time points of the study. There were 329 men who completed all time points: 170 in the UDS group and 159 in the RC group.

As shown in *Figure 4*, the IPSS score increased slightly from the 18-month time point to 5 years. This may have been for a variety of reasons, including stopping long-term treatments or the natural progression of symptoms due to ageing. The trajectory was the same for both groups.

Looking again at the primary outcome results, at 5-year follow-up the results appear to be very similar between the groups; mean IPSS scores of 13.68 and 13.62 for the UD and RC groups respectively. However, non-inferiority could not be confirmed given the wide CI; the difference in means is 0.41 (95% CI −1.10 to 1.93). This is perhaps due to the reduced sample size at 5 years, leading to less power. The point estimate also changed direction between 18 months and 5 years, being in favour of the UD group at 18 months and in favour of RC at 5 years. However, these differences in means, between groups, were very small (−0.33 vs. 0.41). To ensure that this difference was not owing to the change in the sample, the 18-month result for those followed up at 5 years can be found in grey (*Table 6*). This gave a similar result to the Phase I result, suggesting that this shift is a true reflection of symptom patterns over time.

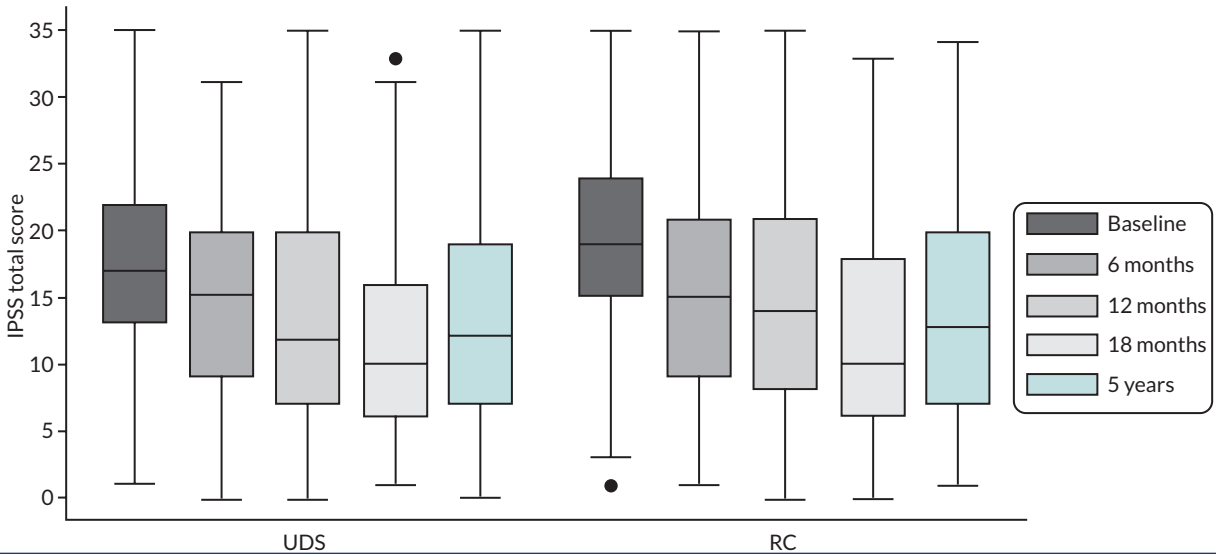

FIGURE 4 International Prostate Symptom Scores over time, for those completing the questionnaire at all time points. Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The figure includes minor additions and formatting changes to the original text.

TABLE 6 Primary outcome of IPSS scores at 5 years

Variable	*N* (U : R)[a]	UDS (*n* = 427)	RC (*n* = 393)	Adj. difference in means[b] (95% CI)
IPSS symptom questionnaire				
Total IPSS score at 5 years (ITT)[c]	208 : 202	13.68 (8.09)	13.62 (8.10)	0.41 (−1.10 to 1.93)
Total IPSS score at 18 months (ITT)[d]	*202 : 196*	*11.72 (7.43)*	*12.63 (7.41)*	*−0.84 (−2.26 to 0.57)*
QoL score[e]	211 : 204	2.67 (1.55)	2.69 (1.52)	−0.01 (−0.31 to 0.28)

R, routine care; U, urodynamics.
a The number of men with available 5-year IPSS data.
b Adjusted for centre and baseline IPSS.
c 4 : 3 men, in the UD and RC groups, respectively, could not be included in the adjusted analysis due to missing baseline IPSS data.
d Assessing the 18-month comparison for those individuals included in the 5-year analysis (6 : 6 men had 5-year data but no 18-month data).
e 2 : 1 men, in the UD and RC groups, respectively, could not be included in the adjusted analysis due to missing baseline IPSS QoL data.

Note
Data are mean (SD) unless otherwise specified.
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

Looking at all men who completed at least one questionnaire, at any time point, and focusing on the means, we can look at the profile plot of the men in the UPSTREAM study (*Figure 5*). At 18 months, the mean scores were 12.61 and 13.10 for the UD and RC groups respectively. At 5-year follow-up, the groups were closer together and higher, with mean scores of 13.68 and 13.62, respectively.

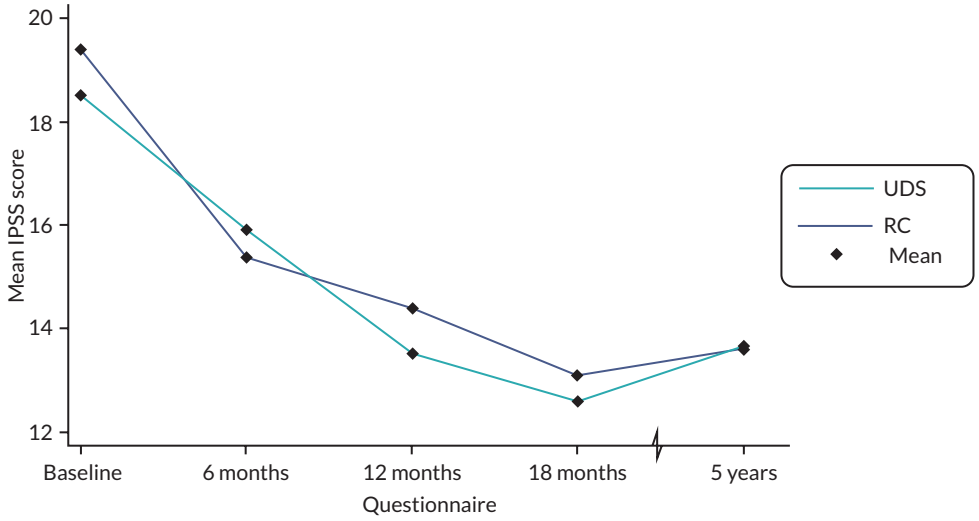

FIGURE 5 Mean IPSS scores over time.

Key secondary outcome

As described in the methods, operational codes were used to obtain the number of LUTS-related surgical procedures from the routine data. Twenty-three codes had been identified as 'surgery related to LUTS' – no additions were made during analysis as all were either already added due to other codes in the spell or not surgical procedures, for example biopsies or scans. Of the 801 individuals who were extracted from routine data, 347 received at least 1 identified related surgical procedure for the treatment of LUTS. This included a proportion of men who underwent introduction of therapeutic substance into the bladder, which is a treatment for storage LUTS used in men with a mixed profile of both

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

storage and voiding LUTS. In total, we identified 454 procedures, 412 in the APC data set and 42 in the OP data set. The most common code was M653 'Endoscopic resection of prostate NEC' with 185 procedures. A full list of codes and frequencies can be found in *Table 7*.

There were eight codes held in reserve, which were deemed 'possible surgery related to LUTS': Y531, Y534, Y535, Z421, Z422, Z424, Z425 and Z429. On inspection, no additions were made to the number of men receiving surgery for LUTS as either they (1) were already included due to other codes in the spell or (2) were not surgical procedures, for example biopsies or scans.

Taking the first procedure, per person, the median time to surgery was 242 days [interquartile range (IQR) 131–435 days]. The minimum number of days between randomisation and surgery was 8 and the maximum was 1812 (5 years).

The procedures picked up by routine data were relatively similar to those captured in Phase I CRFs (*Table 8*). In Phase I, we identified 291 participants who had received surgery and 283 of these were within 548 days (18 months). Based

TABLE 7 Surgical procedures identified in the routine data sets

Code	Meaning	Frequency
M494	Introduction of therapeutic substance into bladder	41
M495	Injection of therapeutic substance into bladder wall	1
M611	Total excision of prostate and capsule of prostate	6
M613	Transvesical prostatectomy	1
M618	Other specified open excision of prostate	1
M642	Implantation of artificial urinary sphincter into outlet of male bladder	1
M651	Endoscopic resection of prostate using electrotome	96
M653	Endoscopic resection of prostate necrotising enterocolitis	185
M654	Endoscopic resection of prostate using laser	33
M655	Endoscopic resection of prostate using vapotrode	1
M662	Endoscopic incision of outlet of male bladder necrotising enterocolitis	26
M668	Other specified other therapeutic endoscopic operations on outlet of male bladder	3
M672	Endoscopic destruction of lesion of prostate necrotising enterocolitis	1
M678	Other specified endoscopic operation on prostate	8
M681	Endoscopic insertion of prostatic stent	1
M682	Endoscopic removal of prostatic stent	1
M683	Endoscopic insertion of prosthesis to compress lobe of prostate	27
M705	Massage of prostate	1
M707	Transurethral radiofrequency needle ablation of prostate	1
M708	Endoscopic insertion of prosthesis to compress lobe of prostate	16
Y089	Unspecified laser therapy to organ NOC	1
Y111	Cauterisation of organ NOC	1
Y171	Electrocauterisation of lesion of organ NOC	1

NOS, not otherwise classified

Note

Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

on appointment and patient availability, 18-month follow-ups were not always within 548 days, hence this slight discrepancy. Of the 283 participants within 18 months in Phase I, the routine data correctly identified 274 (97%). There were nine surgical cases missed by routine data which, on inspection of their codes, did not highlight any additional codes that may have been appropriate to add. The additional eight surgical cases (outside of the 548-day window) were also captured by routine data. So, of the total 291 cases reported in Phase I, 282 matched and 249 had a matching operation date. Looking at those not recorded as receiving surgery in Phase I (n = 499), the routine data correctly identified that 494 (99%) did not receive it in the first 18 months, although 58 then went on to receive it after 18 months. With 282/291 surgical cases being identified and 494/499 non-surgical cases identified, the routine data successfully matched 776/790 (98%) of Phase I findings.

There were 66 additional surgery cases picked up using routine data compared with data collected within the Phase I study (see *Table 8*). The majority of these were after the 18-month follow-up period, although there were seven newly identified individuals with surgery dates within the 18-month period. *Table 9* shows the number of men receiving their first LUTS procedure, in the months and years after randomisation. The majority received surgery within 18 months (slightly delayed in the UD group), and the majority of procedures outside of this window were shortly after it (18–24 months). However, there were still procedures being undertaken in follow-up years 3, 4 and 5, which were relatively balanced between the groups.

Using only the 5-year follow-up routine data, we can see that there were 347 participants who had received surgery for LUTS; 180/415 (43%) in the UD group and 167/386 (43%) in the RC group (*Table 10*). This confirms the result found in Phase I, that there were equal proportions of surgery conducted in each group.

The improvement in IPSS, seen in participants receiving surgery for LUTS, was similar across the two groups, with IPSS scores at 5 years of 11.75 and 11.22 for the UD and RC groups, respectively (*Table 11*). The improvement was much smaller in those not receiving surgery with IPSS scores of 15.33 and 15.65, respectively. Surgery reduced the IPSS score in both groups of the study, but there was no evidence of an interaction with the treatment group, that is no evidence to suggest that surgery decreased the IPSS score more in a particular treatment group (*Table 12*). However, this is an exploratory analysis (not pre-defined), so this result should be seen as hypothesis-generating rather than confirmatory. Exploratory analyses, investigating subgroups where surgery may be more effective, have been carried out previously.[11]

TABLE 8 How well do the Phase I CRF data match the NHS routine data?

	Phase II (NHS data) surgical status			
Phase I (CRF) surgical status (0–18 months)	**No surgery (0–5 years)**	**Surgery conducted (0–18 months)**	**Additional surgery (18 months to 5 years)**	**Total**
Surgery conducted	9	274 (match)	8 (match)[a]	291
No surgery	436 (match)	5 (new)	58 (new)	499
Unknown	9	2 (new)	0	11
Total	454	281	66	801

a Includes eight surgical cases which were outside the 548-day window but also captured in Phase I CRFs (late follow-ups).

Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/.

TABLE 9 Using NHS routine data, when were men receiving their first surgery for LUTS?

Year of follow-up	**Year 1**		**Year 2**		**Year 3**		**Year 4**		**Year 5**		
Months of follow-up	**0–6**	**6–12**	**12–18**	**18–24**	**24–30**	**30–36**	**36–42**	**42–48**	**48–54**	**54–60**	**Total**
Surgery in the UD group	55 (31%)	67 (37%)	24 (13%)	11 (6%)	4 (2%)	5 (3%)	7 (4%)	2 (1%)	3 (2%)	2 (1%)	180
Surgery in the RC group	69 (41%)	49 (29%)	17 (10%)	9 (5%)	8 (5%)	2 (1%)	4 (2%)	5 (3%)	3 (2%)	1 (1%)	167

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 10 Secondary outcome: proportion of men having surgery, difference between groups

Variable	*N* (U : R)[a]	UDS (*n* = 427)	RC (*n* = 393)	OR (95% CI)[b]	*p*-value
Surgery outcome					
Surgery conducted	415 : 386	180 (43%)	167 (43%)	0.96 (0.71 to 1.28)	0.763
No surgery		235 (57%)	219 (57%)		

R, routine care; U, urodynamics.
a There were 802 participants whose details were requested from NHS England. However, we were informed that one participant (in the UD group) could not be traced.
b Logistic regression model, adjusting for centre.

Note
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

TABLE 11 International Prostate Symptom Scores at baseline and 5 years, by surgical status

	UDS Mean (SD)			RC Mean (SD)		
Group	*n* (%)	Baseline	5 years	*n* (%)	Baseline	5 years
IPSS symptom questionnaire						
All participants (ITT)[a]	204	17.98 (6.69)	13.59 (7.99)	199	19.11 (7.27)	13.58 (8.16)
Those who received surgery[b]	99	19.01 (7.03)	11.75 (8.40)	93	21.40 (6.50)	11.22 (8.21)
Those who did not receive surgery[b]	105	17.01 (6.23)	15.33 (7.21)	106	17.09 (7.35)	15.65 (7.56)

a The number of men with available 5-year IPSS data.
b Surgery for LUTS; status according to routine data.

Note
Data are mean (SD) unless otherwise specified.
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/.

TABLE 12 Interaction between surgery and treatment group on IPSS scores

		IPSS at 5 years[b]	Interaction effect[b]	
Variable	*n* (S : NS)[a]	Subgroup-specific difference in means (95% CI)	Difference in means (95% CI)	*p*-value
Subgroup analyses				
Study group				
UDS	99 : 105	−3.25 (−5.35 to −1.14)	2.26 (−0.58 to 5.09)	0.118
RC	93 : 106	−6.30 (−8.57 to −4.03)		

a Surgery (S) : no surgery (NS).
b Adjusted for centre and baseline IPSS.

Sensitivity and subgroup analyses for the primary outcome

All sensitivity analyses led to similar results to the primary outcome (*Table 13*). As the number of men who received UDS was recorded so differently in Phase I and Phase II, this was separated into two separate per-protocol analyses.

To analyse the subgroups, pre-specified formal tests of interaction were employed to explore potential effect modifiers. For the majority of subgroups, there was no evidence to suggest that there were any interactions between

TABLE 13 Sensitivity analyses for the primary outcome of IPSS scores at 5 years

Variable	*N* (U : R)[a]	UDS (*n* = 427)	RC (*n* = 393)	Adj. difference in means[b] (95% CI)
Primary analysis recap				
ITT[c]	208 : 202	13.68 (8.09)	13.62 (8.10)	0.41 (−1.10 to 1.93)
Sensitivity analyses				
Per protocol 1[d]	189 : 190	13.55 (8.16)	13.69 (7.93)	0.36 (−1.19 to 1.91)
Per protocol 2[d]	141 : 164	13.89 (8.30)	13.51 (8.01)	0.63 (−1.27 to 2.53)
As treated analysis 1[f]	201 : 209	13.49 (8.31)	13.80 (7.88)	0.08 (−1.43 to 1.59)
As treated analysis 2[h]	179 : 231	13.93 (8.34)	13.43 (7.90)	0.19 (−1.51 to 1.89)
Repeated measures[h] at 6 months	310 : 272	15.91 (7.44)	15.39 (7.43)	0.97 (−0.20 to 2.14)
at 12 months	300 : 263	13.54 (7.84)	14.40 (7.89)	−0.61 (−1.79 to 0.57)
at 18 months	340 : 329	12.61 (7.92)	13.11 (7.76)	−0.17 (−1.28 to 0.94)
at 5 years	208 : 202	13.68 (8.09)	13.62 (8.10)	0.78 (−0.54 to 2.09)
Multiple imputation[i]	427 : 393	14.23 (9.46)	13.77 (9.24)	0.75 (−0.58 to 2.08)
Adj. for potential confounders[j]	207 : 197	–	–	0.38 (−1.15 to 1.91)
Adj. for time since surgery[k]	208 : 202	–	–	0.54 (−0.92 to 2.00)

R, routine care; U, urodynamics.
a The number of men with available 5-year IPSS data.
b Adjusted for centre and baseline IPSS.
c 4 : 3 men, in the UD and RC groups, respectively, could not be included in the adjusted analysis due to missing baseline IPSS data.
d Removing those who did not comply with their randomised treatment, according to their UDS status in Phase I (3 : 3 excluded from analysis due to missing baseline IPSS data).
e Removing those who did not comply with their randomised treatment, according to their UDS status in Phase II (2 : 2 excluded from analysis due to missing baseline IPSS data).
f According to UDS status, based on Phase I data (3 : 4 excluded from analysis due to missing baseline IPSS data).
g According to UDS status, based on Phase II data (3 : 4 excluded from analysis due to missing baseline IPSS data).
h An interaction term was included in the mixed-effects model; therefore, three coefficients have been presented, one for each time point.
i Multiple imputation by chained equations.
j Additional adjustment for age and presence of comorbidities (4 : 3 excluded from analysis due to missing baseline covariate data).
k Adjusted for the number of days between the surgical procedure and 5-year follow-up time point, a value of 3000 days was imputed where no surgery took place (4 : 3 excluded from analysis due to missing baseline data).

Note
Data are mean (SD) unless otherwise specified.
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

the subgroups on the effect of a study group on IPSS scores or surgery rates at 5-year follow-up (*Tables 14* and *15*). However, there was some strong evidence to suggest that UDS was more effective at reducing IPSS scores in those with higher flow rates. This confirms recent findings, based on Phase I data, that suggest UDS was more effective in this subgroup as UD measures are more strongly predictive of surgical outcome for those with a maximum flow rate > 15.[11] There was also some evidence, albeit weak, to suggest that UDS was more effective in younger men.

Secondary outcomes

Results from the ICIQ-MLUTS and ICIQ-MLUTS-sex questionnaires can be found in *Tables 16* and *17*. Analyses of binary outcomes could not be adjusted for centre due to perfect prediction. Urinary symptoms were very similar between the groups, apart from nocturia which was slightly higher in the RC group (58% vs. 66%), although the difference and

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 14 Subgroup analyses: primary outcome

Variable	IPSS score at 5 years				
	n (U : R)[a]	UDS	RC	Subgroup-specific MD[b] (95% CI)	Interaction effect MD[b] (95% CI); *p*
Subgroup analyses					
Age (continuous)	208 : 202	–	–	–	0.18 (−0.02 to 0.37); 0.068
Flow rate					
≤ 12 ml/second	129 : 126	14.23 (8.40)	12.47 (8.16)	2.20 (0.20 to 4.20)	−4.72 (−7.86 to −1.58); 0.002
> 12 ml/second	72 : 72	13.03 (7.63)	15.50 (7.60)	−2.92 (5.39 to −0.45)	
Maximum voided volume					
< 200 ml	81 : 84	13.80 (8.17)	12.56 (8.37)	0.73 (−1.84 to 3.29)	−1.35 (−4.46 to 1.75); 0.374
≥ 200 ml	120 : 115	13.80 (8.14)	14.31 (7.90)	−0.51 (−2.53 to 1.50)	
Nocturia					
No	51 : 39	11.76 (7.42)	10.82 (7.43)	0.40 (−3.01 to 3.82)	0.02 (−3.66 to 3.71); 0.989
Yes	151 : 159	14.07 (8.23)	14.21 (8.21)	0.39 (−1.44 to 2.22)	
Severity of storage LUTS[c]					
Less substantial (≤ 9)	131 : 111	13.10 (7.50)	12.33 (7.13)	0.46 (−1.31 to 2.24)	−0.43 (−3.56 to 2.70); 0.777
More substantial (> 9)	74 : 89	14.72 (8.97)	15.18 (9.04)	0.16 (−2.89 to 3.22)	

MD, difference in means; R, routine care; U, urodynamics.
a Some men were not included in each analysis due to missing baseline data (0–4 in the UD group and 0–3 in the RC group).
b Linear regression model adjusting for centre and baseline IPSS score.
c The summation of items 2, 4 and 7 in the IPSS questionnaire, split by the median found in Phase I.

Note
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

TABLE 15 Subgroup analyses: key secondary outcome

Variable	Surgery rate				
	n (U : R)	UDS	RC	Subgroup-specific OR[a] (95% CI)	Interaction effect OR[a] (95% CI); *p*
Subgroup analyses					
Age (continuous)	415 : 386	–	–	–	1.01 (0.98 to 1.05); 0.434
Flow rate					
≤ 12 ml/second	249 : 226	131 (53%)	118 (52%)	1.02 (0.71 to 1.46)	1.04 (0.56 to 1.94); 0.892
> 12 ml/second	146 : 142	45 (31%)	42 (30%)	1.06 (0.64 to 1.75)	
Maximum voided volume					
< 200 ml	180 : 168	83 (46%)	77 (46%)	1.01 (0.66 to 1.54)	1.10 (0.62 to 1.95); 0.744
≥ 200 ml	218 : 205	95 (44%)	84 (41%)	1.11 (0.76 to 1.64)	

TABLE 15 Subgroup analyses: key secondary outcome (*continued*)

Variable	Surgery rate				
	n (U : R)	UDS	RC	Subgroup-specific OR[a] (95% CI)	Interaction effect OR[a] (95% CI); *p*
Nocturia					
No	95 : 73	30 (32%)	24 (33%)	0.94 (0.49 to 1.81)	1.14 (0.55 to 2.36); 0.721
Yes	295 : 298	142 (48%)	138 (46%)	1.08 (0.78 to 1.49)	
Severity of storage LUTS[b]					
Less substantial (≤ 9)	231 : 199	97 (42%)	74 (37%)	1.22 (0.83 to 1.80)	0.73 (0.41 to 1.29); 0.277
More substantial (> 9)	169 : 175	81 (48%)	89 (51%)	0.89 (0.58 to 1.36)	

R, routine care; U, urodynamics.
a Logistic regression model not adjusting for centre due to perfect prediction in some subgroups.
b The summation of items 2,4 and 7 in the IPSS questionnaire, split by the median found in Phase I.

Note
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

TABLE 16 Secondary outcome: ICIQ MLUTS scores and items

Variable	*n* (U : R)[a]	UDS Mean (SD)/*n* (%)	RC Mean (SD)/*n* (%)	Difference[b] (95% CI)	*p*-value[b]
ICIQ-MLUTS scores					
ICSmaleVS (voiding scale)[c]	197 : 176	7.06 (4.36)	6.99 (4.33)	0.17 (−0.68 to 1.02)	0.696
ICSmaleIS (incontinence scale)[d]	195 : 173	4.44 (3.57)	4.24 (3.01)	0.24 (−0.39 to 0.87)	0.452
Daytime frequency (> 8 times)	198 : 174	55 (28%)	41 (24%)	1.20 (0.72 to 1.98)	0.484
Nocturia (> 1 times per night)	197 : 172	114 (58%)	114 (66%)	0.68 (0.43 to 1.09)	0.106

R, routine care; U, urodynamics.
a Some men were not included in each analysis due to missing baseline data (6–8 in the UD group and 2–4 in the RC group).
b Linear regression for continuous outcomes (adjusted for centre and baseline scores) or logistic regression for binary outcome (adjusted for baseline scores but not centre, due to perfect prediction).
c Voiding scale, on a scale of 0–20, with larger scores indicating more severe symptoms.
d Incontinence scale, on a scale of 0–24, with larger scores indicating more severe symptoms.

Note
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

p-value did not provide evidence of a difference. Similar percentages were also found in the 18-month follow-up results (59% vs. 67%).[1]

Sexual symptoms were also similar between the groups (*Table 17*), although there was evidence to suggest that the proportion of men reporting painful ejaculation was higher in the RC group than in the UD group (13% vs. 22%). Given that this trend was not present in the other sexual symptom outcomes, this may have been a chance finding.

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 17 Secondary outcome: ICIQ MLUTS-sex items

Variable	*n* (U : R)[a]	UDS Mean (SD)/*n* (%)	RC Mean (SD)/*n* (%)	OR[b] (95% CI)	*p*-value[b]
ICIQ-MLUTS scores					
Erections (reduced or none)	196 : 171	156 (80%)	141 (82%)	0.87 (0.48 to 1.57)	0.647
Ejaculation (reduced or none)	193 : 170	174 (90%)	155 (91%)	1.05 (0.49 to 2.25)	0.892
Painful ejaculation (yes)	177 : 154	23 (13%)	34 (22%)	0.49 (0.25 to 0.95)	0.034
Urinary symptoms affected sex life?	190 : 170	146 (77%)	124 (73%)	1.01 (0.60 to 1.75)	0.959

R, routine care; U, urodynamics.
a Some men were not included in each analysis due to missing baseline data (7–10 in the UD group and 3–6 in the RC group).
b Logistic regression, adjusted for baseline scores but not centre, due to perfect prediction.

Note
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

Mortality, according to Office of National Statistics records

The mortality data were extracted from ONS death records. At 5-year follow-up, there had been a total of 28 deaths in the UD group and 26 deaths in the RC group (*Table 18*).

As observed in the initial phase of the study,[12] there were more deaths in the UD group than in the RC group. However, the number of deaths in each group had evened out by the end of the 5-year follow-up: 28 versus 26 in the UD and RC groups, respectively. The causes of death in each Phase can be found in *Appendix 2*, *Tables 24* and *Appendix 3*, *Table 25*.

TABLE 18 Mortality, according to ONS death records

	UDS[a] (*n* = 415)	RC[a] (*n* = 386)
Mortality		
Death during Phase I	11	3
Death during Phase II	17	23
Median time to death (IQR) in months	35 (15–50)	40 (30–52)

a There were 802 participants whose details were requested from NHS England. However, we were informed that one participant (in the UD group) could not be traced.

Chapter 4 Economic evaluation

Some of the text in this chapter is reproduced with permission from Clout *et al*.[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Introduction

Aim and objectives for economic evaluation

The aim of the economic evaluation was to understand whether a care pathway for treating bothersome LUTS which includes UD testing is cost-effective 5 years post randomisation, compared to a care pathway that omits UD testing.

Primary objective

The primary objective of the within-study economic evaluation was to estimate, from an NHS secondary care perspective, the 5-year post randomisation cost-effectiveness of the UD pathway to treat men with bothersome LUTS versus the RC pathway.

Secondary objective

The secondary objective was to report the differential use of NHS resources for LUTS-related diagnostic tests and therapy. For example, this included additional diagnostic testing (e.g. UDS) and/or the decision to perform surgery.

Methods

Health economic analysis plan

A HEAP was developed prospectively by the study team and made available on an open-access repository (Pure).[13] The HEAP was in line with NICE's reference case[14] and was approved by the TSC prior to the analysis and followed best practices on what content should be included in a HEAP.[15] The HEAP referred to the further follow-up of existing UPSTREAM Phase I participants, specifically at 5 years post randomisation (UPSTREAM Phase II).

Population

Characteristics of the study population and setting are provided in *Chapter 3* for the clinical outcome. In brief, the economic analysis followed the ITT principle and analysed all 820 participants according to the care pathway they were randomised to at the start of Phase I of the study. The study was conducted in the UK where the health system is publicly funded and free at the point of access.

Intervention and comparator

Descriptions of the intervention and comparator groups, as well as the methods for allocating participants to study groups, are detailed in the Phase I report[12] and Phase I Protocol (version 4, 29 September 2016). No new study groups were introduced for Phase II. The care pathway for the assessment of men with LUTS without UDS (RC) may include the following diagnostic tests: physical examination of the prostate, symptom score measurement, a bladder diary and flow rate testing with PVR scan.[16]

Study perspective and time horizon

An NHS secondary care perspective excluding accident and emergency (A&E) resource was used. The primary care and A&E resource use had been similar between the two groups at 18 months, and there was no clinical reason why this should change after 18 months. All analyses for Phase II were conducted over a time horizon of 5 years post randomisation. Five years was deemed long enough to capture any important differences in costs and outcomes between the two groups.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Measurement of resources

All resources involved in the two care pathways (UDS and RC) were measured using the HES APC and HES OP data sets. Access and data linkage for these data sets are described in *Chapter 2* for the clinical outcome. These patient-level data sets captured resource use for: admitted day case and overnight care, admitted emergency care for non-elective and elective visits, and OP consultations and procedures.

All care that could have been related to bothersome LUTS were identified in the HES APC data sets using a list of codes for the following fields: treatment specialty, procedures and diagnoses. As no LUTS-specific code lists were identified on the Health Data Research UK Phenotype Library, we developed our own code list. This list of codes was devised by a clinician with expertise in urology (MD). The procedural codes included in the data sets were called the Classification of Interventions and Procedures version 4 (OPCS-4) codes. For HES APC records that represented care which did not involve any procedures, primary diagnosis codes [*International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision (ICD-10)] were checked by a clinician (MD). HES OP records which represented OP consultations but did not have a procedural code could not be filtered by primary diagnosis code (ICD-10) because these data are poorly coded in the HES OP data set. Consequently, all OP consultations were included in the primary analysis.

Valuation of resources

Healthcare Resource Groups (HRGs) were used to enable the valuation of the HES APC and HES OP resource data. HRGs are the classification system used in the NHS in England to group patient care activity into clinically meaningful groupings based on expected clinical resource use. They provide a national standardised way to cost clinical care. In UPSTREAM Phase II, the procedure (OPCS-4) and diagnosis (ICD-10) codes reported in the patients' records were used to group the patients' healthcare activities into HRG codes. The HES APC and OP care resource use data were processed and grouped into HRG codes using the latest version of the NHS Reference Costs Grouper which uses the enhanced HRG4+ classification system.[17] All records were assigned a core HRG code to represent the primary reason for the admission. In addition, some records were assigned one or more 'unbundled' HRG codes to represent a situation where an admission resulted in a significant element of cost or activity (e.g. diagnostic imaging, high-cost drugs). We used the latest available NHS reference costs (2020–1) to assign a unit cost to all the core and unbundled HRG codes in our data set.[18] Tables within the reference costs used for admitted patient care (APC) unit costs included: Elective Inpatients, Non-Elective Inpatients, Non-Elective Short Stay, Day Case and Diagnostic Imaging. Tables for OP care included: Consultant Led, Non-Consultant Led, Outpatient Procedures and Diagnostic Imaging. Unit costs for the core HRGs were assigned at the finished consultant episode (FCE) level, rather than spell level. This was because the NHS reference costs were based on FCEs.[19] FCEs for unbundled HRG codes were initially costed separately and then added to the total aggregate cost for each core HRG.

Measurement of outcomes

The primary economic outcome was QALYs derived from health-related quality of life scores, obtained using the EQ-5D-5L questionnaire. QALYs and the EQ-5D-5L questionnaire are recommended by NICE.[20] In Phase I, participants completed the EQ-5D-5L questionnaire at baseline, 6, 12 and 18 months. In Phase II, the EQ-5D-5L questionnaire was included in the '5-year follow-up' participant self-report questionnaire booklet. Participants were able to complete the questionnaire via post, online or telephone.

Valuation of outcomes

Participant's EQ-5D-5L health states for the multiple time points (baseline, 6, 12 and 18 months, and 5 years) were mapped to the published UK population EuroQol-5 Dimensions, three-level version valuation set using a validated mapping tool by van Hout *et al.*[21] This mapping tool was the recommended approach by NICE during Phase I and so the same tool was applied for consistency. QALYs were constructed using the area under the curve approach, which involved combining the individual participant's EQ-5D-5L index values with their survival data. For participants who died during the 5-year period, zero was imputed for their utility value following their date of death. Survival data were extracted from the HES-ONS-linked mortality data (as described in the *Chapter 2* section for the clinical outcome).

Analysis

For the primary analysis, both costs and QALYs falling in years 2, 3, 4 and 5 were discounted at 3.5% in accordance with guidance from NICE. Stata version 17.0 was used for all health economic analyses.

Summary statistics

Mean resource use and QALYs were estimated and reported by the study group. SD and the number of participants included in each time point were also presented.

Incremental adjusted mean costs and utility for the 5 years were estimated using multiple imputation with chained equations and predictive mean matching (STATA function: mi impute chain pmm). More specifically, we created 60 imputed data sets by building regression models that used all variables included in the analysis. HES data had the same amount of missingness for the 5-year follow-up period. EQ-5D-5L utility scores had different patterns of missingness over the follow-up time points, therefore utility scores were imputed at each time point (6, 12 and 18 months, and 5 years), and then aggregated. Please see *Appendix 4* for the multiple imputation code. The STATA function mi passive was used to run seemingly unrelated regression (SUR) on the multiple imputed data sets. SUR allows for the joint modelling of cost and QALY regressions and therefore can account for the potential correlations between the cost and QALY differences. Both the costs and QALY estimates were adjusted for centre and baseline IPSS. Additionally, the QALY estimates were adjusted for baseline utility. The resulting estimates were combined through the net benefit framework to estimate the incremental net monetary benefit (INMB) statistic. The incremental cost-effectiveness ratio was not created because the intervention group was dominant.

Characterising sampling uncertainty

Uncertainty in our INMB estimate due to our sampling variation was explored and represented using CIs and cost-effectiveness acceptability curves (CEACs). More specifically, the CEACs presented the probability of the UD group being cost-effective compared to RC for a range of willingness-to-pay (WTP) thresholds, including NICE's recommended cost-effectiveness thresholds of £20,000–30,000 per QALY.

Characterising methodological uncertainty

Uncertainty in the methodological assumptions made for the present economic evaluation was assessed through a number of sensitivity analyses. This involved making plausible changes to key methodological assumptions in order to understand how changes in assumptions impacted the cost-effectiveness result. Sensitivity analyses included: varying the discount rate to 1.5%, an alternative discount rate used by NICE, and reducing OP consultation costs in both groups by 62% because information about a patient's diagnosis is not mandated as part of the Outpatient Commissioning Data Set from which the HES OP data set is derived. This meant it was not possible to indicate whether the OP consultations were likely to be LUTS-related. Nonetheless, it was possible to include only consultations that took place in LUTS-related treatment specialties. A reduction of 62% was chosen as this reflected the proportion of records that were not LUTS-related in the inpatient (APC) data set, after filtering by treatment specialty, procedure and diagnosis codes.

Results

Summary of main results

The study had high rates (97.7%, *n* = 801) of HES and survival data from the electronic health records. Just 12 participants in the UD intervention group and 7 participants in the RC group were missing HES and survival information.

Similarly, baseline completion rates for the EQ-5D-5L self-report questionnaire were high for both the UD intervention group (91.6%, *n* = 391) and the RC group (93.4%, *n* = 393). Completion rates for the EQ-5D-5L questionnaire declined over time due to deaths, no consent provided for the 5-year follow-up or attrition due to the long follow-up period. *Table 19* shows that completion rates and EQ-5D-5L scores were similar across both groups. The lower score observed in both groups at the 5-year follow-up may reflect the fact that most deaths (74.1%, *n* = 40) took place between the 18-month and 5-year follow-up period. Excluding the 54 participants for whom an EQ-5D-5L score of zero was assigned following their death, the EQ-5D-5L questionnaire was completed at all five time points for just over a third (*n* = 317, 38.66%) of participants.

National Health Service secondary care resource use data were reviewed for 801 participants. For the 19 participants who had no resource use data: 11 men requested their electronic health records were not accessed; 7 men withdrew their consent during the first 18-month period when they completely withdrew from the study; NHS England were unable to link 1 participant from the UD group, despite being provided with key data linkage identifiers (date of birth and NHS number).

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

TABLE 19 Mean unadjusted EQ-5D-5L scores and completion rates at each data collection time point

	UDS (*n* = 427)		RC (*n* = 393)	
Time point	*n*[a] (Completion rate)	Mean (SD)	*n*[a] (Completion rate)	Mean (SD)
Baseline	391 (91.6%)	0.782 (0.217)	367 (93.4%)	0.777 (0.233)
6 months[b]	311 (72.8%)	0.786 (0.224)	271 (69.0%)	0.770 (0.255)
12 months[c]	303 (70.9%)	0.765 (0.257)	265 (67.4%)	0.787 (0.221)
18 months[d]	324 (75.9%)	0.754 (0.271)	312 (79.4%)	0.793 (0.241)
5 years[e]	250 (58.5%)	0.644 (0.357)	237 (60.3%)	0.644 (0.340)

a Available case.
b Including three patients who died between the 0 and 6 months.
c Including four additional patients who died between 6 and 12 months.
d Including seven additional patients who died between 12 and 18 months.
e Including 40 additional patients who died between 18 months and 5 years.

For the HES APC data set, 5051 records referred to a FCE that took place during the 5-year follow-up period. After filtering by treatment specialty, procedure and diagnosis codes, and after removing duplicates, just 880 FCE records were identified as being relevant for the primary analysis. For the HES OP data set, 800 of the 801 participants had 25,108 records that referred to a FCE that took place during the 5-year follow-up period. After filtering by treatment specialty, a total of 9467 FCE records remained. Nearly all of these records (94%, *n* = 8899) were coded as having an unknown diagnosis and so it was not possible to filter the HES OP data by diagnosis codes. Similarly, only a third (28%, *n* = 2651) of records had procedure codes. After filtering by procedure codes if a record had any, and after removing duplicates and cancelled OP appointments, a total of 6102 HES OP records remained for the primary analysis.

Table 20 shows the mean NHS secondary care resource use for the 97.7% (*n* = 801) of participants for whom it was possible to collect HES APC and HES OP data. Overall, *Table 20* shows that there were similar amounts of secondary care resource use across both groups over the 5-year period.

The most common type of APC was TURP surgery. TURP surgery rates were similar in both groups. By contrast, UD testing was twice as likely to be provided to participants admitted to hospital in the UD intervention group compared to RC group. Similarly, in OP care, UD testing was around a quarter more likely in the UD intervention group compared to the RC group. These higher rates observed in the UD group were expected given the nature of the UD intervention. The mean resource use for other admitted and OP procedures were low and similar across both groups.

On average, participants had around one initial face-to-face OP consultation and around four follow-up face-to-face consultations in a treatment specialty in which a LUTS-related consultation could have occurred (see *Table 20*).

TABLE 20 Mean resource use for each group for participants with complete resource use data (*n* = 801)

		UDS (*n* = 415[a])	RC (*n* = 386[a])
HRG4 + code	HRG4 + code	Resource use mean (SD)	Resource use mean (SD)
APC: LUTS-related surgical or UD procedures			
Intermediate Endoscopic Bladder Procedures	LB14Z	0	0.003 (0.051)
Introduction of Therapeutic Substance into Bladder	LB17Z	0	0.005 (0.102)
Major Open, Prostate or Bladder Neck Procedures with CC Score 2+	LB21A	0.005 (0.069)	0
Major Open, Prostate or Bladder Neck Procedures with CC Score 0–1	LB21B	0	0.005 (0.072)
Transurethral Prostate Resection Procedures with CC Score 6+	LB25D	0.012 (0.109)	0.013 (0.134)

TABLE 20 Mean resource use for each group for participants with complete resource use data (*n* = 801) (*continued*)

HRG4 + code	HRG4 + code	UDS (*n* = 415[a]) Resource use mean (SD)	RC (*n* = 386[a]) Resource use mean (SD)
Transurethral Prostate Resection Procedures with CC Score 3–5	LB25E	0.065 (0.247)	0.042 (0.200)
Transurethral Prostate Resection Procedures with CC Score 0–2	LB25F	0.299 (0.469)	0.332 (0.509)
Intermediate Endoscopic, Prostate or Bladder Neck Procedures, CC Score 2+	LB26A	0.010 (0.098)	0.003 (0.051)
Intermediate Endoscopic, Prostate or Bladder Neck Procedures, CC Score 0–1	LB26B	0.012 (0.109)	0.005 (0.072)
Dynamic Studies of Urinary Tract	LB42A	0.133 (0.386)	0.070 (0.284)
Implantation of Artificial Urinary Sphincter	LB50Z	0.002 (0.049)	0
Major Robotic, Prostate or Bladder Neck Procedures	LB69Z	0.010 (0.098)	0.003 (0.051)
Complex Endoscopic, Prostate or Bladder Neck Procedures, CC Score 2+	LB70C	0.012 (0.109)	0.013 (0.113)
Complex Endoscopic, Prostate or Bladder Neck Procedures, CC Score 0–1	LB70D	0.046 (0.221)	0.041 (0.200)
Minor Prostate or Bladder Neck Procedures	LB78Z	0.010 (0.098)	0.016 (0.124)
APC: other LUTS-related care			
Other LUTS-related care	Varies	0.465 (1.285)	0.508 (1.415)
OP care: LUTS-related surgical or UD procedures			
Intermediate Endoscopic Ureter Procedures	LB09D	0	0.003 (0.051)
Intermediate Endoscopic Bladder Procedures	LB14Z	0.014 (0.170)	0.008 (0.088)
Minor Bladder Procedures	LB15E	0.125 (0.353)	0.111 (0.346)
Introduction of Therapeutic Substance into Bladder	LB17Z	0	0.083 (1.260)
Attention to Suprapubic Bladder Catheter	LB18Z	0.002 (0.049)	0
Intermediate Endoscopic, Prostate or Bladder Neck Procedures, CC Score 0–1	LB26B	0	0.003 (0.051)
Dynamic Studies of Urinary Tract	LB42A	1.043 (1.148)	0.793 (1.186)
Minor or Intermediate, Urethra Procedures	LB55A	0.002 (0.049)	0.008 (0.114)
Diagnostic Flexible Cystoscopy	LB72A	0.347 (0.808)	0.350 (0.856)
Minor Prostate or Bladder Neck Procedures	LB78Z	0.007 (0.085)	0
Percutaneous Ablation of Lesion of Prostate	YL30Z	0.002 (0.049)	0
OP consultations in LUTS-related treatment specialties[b]			
Face-to-Face Attendance, First	WF01B	1.214 (1.456)	1.251 (1.647)
Face-to-Face Attendance, Follow-up	WF01A	4.337 (4.307)	4.490 (4.700)
Non-Face-to-Face Attendance, First	WF01D	0.058 (0.272)	0.088 (0.326)
Non-Face-to-Face Attendance, Follow-up	WF01C	0.393 (1.345)	0.329 (1.053)
Multiprofessional Face-to-Face Attendance, First	WF02B	0.002 (0.049)	0
Multiprofessional Face-to-Face Attendance, Follow-up	WF02A	0.005 (0.069)	0.005 (0.072)

HRG4 +, Healthcare Resource Groups version 4; CC, complication and comorbidity level where a higher CC score indicates the procedure was for a patient with greater severity and/or complexity, and it is expected to result in additional resource use and costs.

a Medical records available for 801/820 participants.

b LUTS-related specialties included: General Surgery Service, Urology Service, Anaesthetic Service, Endocrinology Service, Diabetes Service, Renal Medicine Service, Medical Oncology Service, Neurology Service, Clinical Neurophysiology Service and Clinical Oncology Service.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Table 21 shows both groups incurred the majority of their total secondary care costs in the first and second year of the study. The UD intervention group incurred £136 and £168 more per person than the RC group during the first and second year, respectively. These costs could partly be explained by the extra UDS test added to the care pathway for those in the UD intervention group.

Table 20 shows that adjusted mean costs reduced over time for both groups after the first 2 years. The fifth-year follow-up period fell between October 2018 and December 2021. This was during the COVID-19 pandemic. However, there is no evidence to indicate the COVID-19 pandemic impacted the 2 groups nor the 26 participating urology departments differently. During the fifth year of the study, the UD intervention group resulted in lower mean costs (−£138) compared to RC.

Table 21 reports the adjusted mean QALYs for each of the 5 years. The table shows that by the end of year 1, the UD group had gained slightly more QALYs than the RC group. For the subsequent 4 years, it was the RC group which gained slightly more QALYs than the UD group. QALYs gained during each year decreased over time which may partly reflect the observation that at least 14.8% (*n* = 54) of the participants for whom we had survival data, died over the 5 years.

Cost-effectiveness analysis

Table 22 shows that over the 5-year time horizon, adjusted incremental mean costs were slightly higher (£176.63, 95% CI −£464.06 to £817.32) in the UD intervention group compared to RC. Similarly, adjusted incremental mean QALYs were slightly lower (−0.039, 95% CI −0.152 to 0.073) for the UD intervention group. The INMB statistic was negative at −£962.62 (95% CI −£3323.54 to £1398.30). This indicates the UD group is unlikely to be cost-effective when applying the UK's recommended threshold of £20,000 per QALY. For all estimates reported in *Table 22*, the 95% CIs were wide and crossed zero which indicated there is uncertainty in our results.

TABLE 21 Mean and incremental costs and QALYs presented by year[a]

			Adjusted mean (95% CI)[a]		Incremental adjusted mean (95% CI)[a]	
	Study group	*n*	Costs (£)	QALYs	Costs (£)	QALYs
Year 1	UDS	393	£2449.31 (£2189.01 to £2709.61)	0.772 (0.759 to 0.784)	£135.86 (−£242.50 to £514.23)	0.003 (−0.015 to 0.021)
	RC	427	£2313.44 (£2041.74 to £2585.15)	0.769 (0.755 to 0.782)		
Year 2[b]	UDS	393	£986.37 (£794.74 to £1178.01)	0.737 (0.721 to 0.754)	£167.77 (−£110.12 to £445.66)	−0.019 (−0.042 to 0.005)
	RC	427	£818.61 (£619.58 to £1017.64)	0.756 (0.739 to 0.773)		
Year 3[b]	UDS	393	£424.55 (£294.65 to £554.46)	0.691 (0.673 to 0.709)	£9.21 (−£179.82 to £198.24)	−0.013 (−0.039 to 0.014)
	RC	427	£415.34 (£279.52 to £551.16)	0.704 (0.685 to 0.722)		
Year 4[b]	UDS	393	£368.37 (£256.29 to £480.46)	0.662 (0.643 to 0.681)	£1.55 (−£160.87 to £163.96)	−0.006 (−0.152 to 0.073)
	RC	427	£366.83 (£250.44 to £483.22)	0.668 (0.649 to 0.687)		
Year 5	UDS	393	£235.80 (£118.19 to £353.41)	0.628 (0.609 to 0.648)	−£137.76 (−£309.32 to £33.81)	−0.004 (−0.033 to 0.024)
	RC	427	£373.56 (£250.03 to £497.09)	0.633 (0.613 to 0.653)		

a Adjusted for baseline IPSS score and centre for costs. In addition, QALYs were adjusted for baseline utility. SUR and multiple imputation were used to estimate the values.
b Utilities at years 2, 3 and 4 were linearly interpolated.

Uncertainty due to our sampling variation was explored and represented using a CEAC (*Figure 6*). The CEAC shows the proportion of INMB estimates that are positive at each given WTP threshold given on the *x*-axis. At a WTP threshold of £20,000, the INMB estimate is likely to be negative with a 21% probability of being cost-effective. Lastly, our sensitivity analyses are presented in *Table 22*. The slightly higher mean costs and lower QALYs in the UD group, and the negative INMB estimates along with the wide 95% CIs, align with the findings from our main analysis. That is to say, the sensitivity analyses did not change how we interpreted the main findings (see *Table 22*).

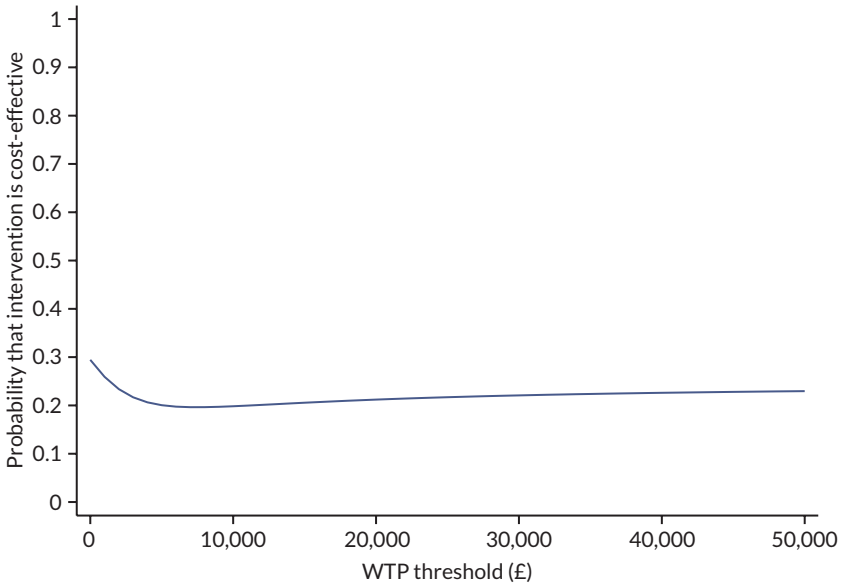

FIGURE 6 Cost-effectiveness acceptability curve for primary analysis.

TABLE 22 Primary analysis and sensitivity analyses

		Adjusted mean (95% CI)[a]		Incremental adjusted mean (95% CI)[a]		
Study group	***n***	**Costs (£)**	**QALYs**	**Costs (£)**	**QALYs**	**INMB (£) at £20,000/QALY (95% CI)**
Primary analysis: multiple imputation with costs and QALYs discounted at 3.5% per year						
UDS	393	£4464.41 (£4024.09 to £4904.73)	3.490 (3.413 to 3.567)	£176.63 (−£464.06 to £817.32)	−0.039 (−0.152 to 0.073)	−£962.62 (−£3323.54 to £1398.30)
RC	427	£4287.78 (£3827.45 to £4748.11)	3.529 (3.450 to 3.608)			
Sensitivity analysis: multiple imputation with costs and QALYs discounted at 1.5% per year						
Intervention	393	£4542.09 (£4090.00 to £4994.19)	3.616 (3.535 to 3.697)	£169.21 (−£488.61 to £827.04)	−0.041 (−0.158 to 0.077)	−£983.33 (−£3446.02 to 1479.35)
RC	427	£4372.88 (£3900.23 to £4845.53)	3.656 (3.574 to 3.739)			
Sensitivity analysis: multiple imputation reducing OP consultation costs by 62% per person based on the proportion of care that was not LUTS-related in the APC data set after filtering by treatment specialty						
Intervention	393	£3854.66 (£3433.51 to £4275.80)	3.489 (3.413 to 3.566)	£185.87 (−£421.74 to £793.48)	−0.047 (−0.156 to 0.0621)	−£1127.15 (−£3414.44 to £1160.14)
RC	427	£3668.79 (£3232.37 to 4105.20)	3.536 (3.458 to 3.615)			

a Adjusted for baseline IPSS score and centre. In addition, QALYs were adjusted for baseline utility.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Discussion

Main findings

Our analyses indicate that, over a 5-year time horizon, the UD group is unlikely to be cost-effective, a result of the UD group having slightly higher costs (£177) and slightly lower (−0.039) QALYs compared to RC. A breakdown of the results by year indicates that the difference in costs was most distinguished during the first 2 years. As resource use was similar across groups for all admitted and OP care categories, it is possible that this difference in costs during the initial couple of years may be driven by the additional costs associated with the extra test in the UD group. The average unit cost for UD testing in 2020–1 was £268.

By the third and fourth year of the study, the difference in costs reduced to almost zero. Furthermore, by year 5, the results had changed direction whereby participants in the UD group were on average −£138 less costly than RC. Nonetheless, this difference in resource use in year 5 was driven by a small number of patients. In particular, there was one patient who had high resource use in the RC group. In terms of QALYs, it is not clear why the UD group gained fewer QALYs after the initial year. Missing resource use data and survival rates over the 5-year period were similar across groups. As costs and QALYs in the two groups were similar over the 5 years, these results should be interpreted alongside the primary outcome. Sources of uncertainty in our analyses were partly due to our sampling variation as characterised in the CEAC and 95% CIs. Our sensitivity analyses did not change how we judged our main results.

Strengths and limitations

A key strength of our study is that for the full 5-year follow-up period we had very high rates (97.7%, *n* = 801) of complete resource use and survival data. Furthermore, for our APC analysis, we identified and limited our analysis to 880 records that were based on a LUTS-related clinical treatment specialty, procedure and diagnosis codes. A cost list was developed by a urology clinician and was effectively applied independently by two researchers. This is a strength of our study as APC is a main cost driver for secondary care. By identifying and incorporating only LUTS-related APC costs, we were able to assign unit costs at a level of detail that accounted for the admission type (e.g. elective, non-elective, day case) and treatment specialty.

A limitation of our study was that it was not possible to know which OP consultations were LUTS-related due to diagnosis codes not being mandated as part of the Outpatient Commissioning Data Set which our HES OP data set derives from. We therefore included all OP consultations for both groups and explored this methodological decision in a sensitivity analysis. Another limitation of our study is our NHS perspective excludes primary care and A&E attendance. Primary care and A&E resource use had been similar between the two groups at 18 months, and there was no clinical reason why this should change after 18 months. Nevertheless, we did consider capturing A&E attendances using the Emergency Care Data Set (ECDS). However, given the time span of the study, we were advised we would need to use both the ECDS and HES A&E data sets. This would incur substantial extra costs and analytical time. In Phase I, there had only been a cost difference between the study groups of A&E attendances of £9 at 18 months; therefore, a decision was made to only include A&E use from the OP data set. However, we decided not to include this information in the analysis because of the substantial under-reporting of its use [1.9% (*n* = 8/415: UD group) and 1.8% (*n* = 7/386: RC)]. Lastly, we compared just two care pathways in England which limits the generalisability of our findings.

In terms of missing outcome data, our survival data enabled us to identify participants who had missing EQ-5D-5L data due to dying during the 5-year period. Other reasons for missing EQ-5D-5L data at each time point were unknown, although the 5-year follow-up time point experienced the largest proportion of missing data which is likely to be due to the long follow-up period.

Chapter 5 Interpretation

The UPSTREAM study (Phase I) identified that a care pathway including UDS was non-inferior to RC in terms of IPSS at 18 months but with no reduction in the rate of surgical treatments. Hence, UD should not be offered routinely in men considering surgery for LUTS. The 18-month time frame of the study did not encompass all diagnostic tests and interventions expected within the pathway. This carried the risk of differentially affecting the two groups because the UD group underwent an additional diagnostic test. 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 after 5 years in the study was underpowered to confirm non-inferiority. This strongly backed up the credibility of the study's main findings.

Response rates to requests to complete questionnaires, including by IPSS and EQ-5D-5L, were reasonable. The HES routine data appeared to give reliable information regarding surgical interventions, thus providing data for a significant majority of the men. It could not, however, achieve the same for diagnostic testing, in particular UDS. This appeared to be due to inconsistent coding within the data. Hence, it was not possible to ascertain a comprehensive description of invasive testing after the 18-month window of the UPSTREAM Phase I study. Likewise, HES data about consultations did not incorporate reasons for consultation, limiting the interpretation of the OP data.

The study has considerable importance in the management of MLUTS. There has long been a split in attitudes of departments towards invasive testing to identify the presence of BOO. Many units made referrals to undertake such testing in all cases, while others avoided it or made limited use of it. UPSTREAM clearly identifies that invasive testing did not reduce surgery rates (key secondary outcome) and that symptom outcomes were 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 symptom outcome was bad (failure to improve or indeed deterioration in IPSS). Post hoc analysis shows the subgroups of men likely to see 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 adverse events associated.

The main strength of the study relates to the ability to identify the key outcome of surgery from HES information 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 morbidities, meant that fewer men were available to complete patient-reported outcomes at 5 years than had been at 18 months. No other high-quality study reports 5-year outcomes from a randomised trial of UDS in MLUTS. The study population was less diverse than the general population, so the conclusions may not be fully generalisable in all contexts.

A major challenge faced was the completion of the procedures required to obtain HES data, which were laborious. This reflects the potential sensitivity of healthcare information, so incorporates a necessary security aspect. It necessitated a study extension to cope with the long timescales to get permissions and receive the data. Recovery from the COVID pandemic influenced the performance of the teams responsible for providing the data. This has strengthened awareness of requirements for such information within the BTC and provides the methodological opportunity for future research.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Chapter 6 Patient and public involvement

During the funding application and development of the UPSTREAM Phase I study, a patient panel of eight volunteers from the North Bristol NHS Trust research and innovation department provided insight and feedback. In addition, 15 men were consulted at flow rate/UDS clinics at 2 hospitals about the procedures that would, and could, occur during the study. All patient representatives regarded invasive UDS as acceptable and this pre-study feedback from patient and public representatives helped develop the initial study design and participant-facing documents. Throughout Phase I, we sought insight on ways to improve the study experience for men, resulting in changes to documentation and processes.

UPSTREAM Phase II also included patient and public involvement (PPI) throughout its lifecycle, via mixed methods (e.g. remote communication and face-to-face meetings).

During the development of the additional funding request for the follow-up phase, members of the existing PPI group from Phase I provided guidance and feedback about what the follow-up study should focus on and how to achieve this. Members of this panel also contributed to the design of the letter of invitation to participate, and the PIS, as well as other patient-facing materials (e.g. newsletters).

Throughout the study, there has been a continued presence of two patient representatives at the PMG and TSC meetings. They have advised on study progress, methods to improve retention, content and design of newsletters and continued development of the project. As per our dissemination policy, methods of communicating the results will include plain English summaries on websites and other more accessible forms for patients. We are working with our PPI panel to define the best methods to disseminate results to patients, including interaction with relevant charities.

Chapter 7 Equality, diversity and inclusion

Inclusion criteria for UPSTREAM Phase I were wide, with all men seeking treatment for bothersome urinary symptoms being considered for participation. Exclusion criteria were mostly clinical and related to urological concerns; however, men who were unable to complete outcome assessments and those with relevant neurological disease (e.g. stroke) were not considered to be eligible. Neurological disease affects the disease process significantly, and it is clinically considered to necessitate routine use of UDS for diagnostic certainty. This may have resulted in the exclusion of some groups of patients.

For Phase II, all Phase I participants were considered for inclusion. Patients were not included in Phase II if: (1) they had withdrawn from Phase I, (2) they had died or had serious ill health (as identified by site teams), (3) they had declined further study contact, or (4) they had declined consent to further follow-up/left this blank on the consent form. Baseline characteristics for those included and not included in the 5-year questionnaire analysis have been reported in *Table 3*. There do not appear to be any substantial differences between the two groups, and it is not anticipated that any particular patient groups have been excluded between the two phases. Phase II participants were mainly of white ethnicity (92%), slightly higher than the average in randomised controlled trials of 86%.[22] For future studies, it may be important to consider methods to encourage participation of black men, due to their under-representation within this research. Participants were spread quite evenly across the Index of Multiple Deprivation (IMD) score quintiles [Quintile 1 (most deprived) = 11%, Quintile 2 = 16%, Quintile 3 = 25%, Quintile 4 = 23%, Quintile 5 (least deprived) = 26%]. The majority of participants (64%) had comorbidities at baseline.

The full spectrum of potential urinary symptoms was reported by men in UPSTREAM Phase I, including voiding symptoms, storage symptoms and nocturia. This reflects the real-life experience of urology departments and shows the study had a representative population. Participants were recruited from 26 hospital sites across England and, as described above, were quite evenly distributed across deprivation scores. Patient involvement is described above. Patient representatives were identified from the same setting as UPSTREAM participants, that is urology clinics in hospitals in England.

Members of the UPSTREAM PPI group were involved in the design of participant documents, including letters of invitation, information sheets and newsletters. Through this involvement, it is hoped that materials will be inclusive and accessible to the target study population.

The research team was representative of a variety of groups, experiences and expertise.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Chapter 8 Impact and learning

The assessment of MLUTS is an important part of service delivery within urology. Surgery to relieve BOO is one of the major activities for healthcare systems. The importance of achieving good symptom outcomes is a big priority for any patient undergoing such surgery, and it may well influence their willingness to consider having the operation. Phase I of the UPSTREAM study identified that symptom outcomes are non-inferior where UDS is included in the assessment pathway while Phase II, described in this report, identified very little difference between the two groups at 5 years. The rates of surgery in the two groups were similar at both 18 months and 5 years following randomisation. Accordingly, urology departments are now in a position to resolve a long-standing dichotomy between routine use of UDS for all men considering such surgery versus selective use, in favour of the latter. This will bring advantages for service delivery in reducing the number of steps in the assessment pathway, thereby shortening the time to treatment.

Alongside the primary outcome of the study, additional post hoc analyses provide significant learning opportunities for urology departments. Firstly, the research group has reported which baseline characteristics are relevant when deciding whether to investigate further with UD testing.[23] Secondly, the importance of quality assurance was underlined when data submitted to the study were analysed alongside another NIHR HTA study looking at issues of equipment maintenance and interpretation of findings in both UDS and flow rate testing.[24] In addition, the importance of good quality, the broad PROMs was identified, notably the relevance to patients of key symptoms like post-micturition dribble and incontinence.[25] Urology departments also need to consider the opinions expressed in the qualitative research.[6,26]

Chapter 9 Implications for decision-makers

The clinical findings that symptom outcomes are non-inferior at 18 months, though surgery rates are not reduced, strongly indicate that UDS should not be used routinely for all patients. Selective use is still needed, given the identification of patients who had a bad outcome from surgery. The data can be used to identify baseline characteristics of patients with the best chance of symptomatic improvement, along with those at greater risk of a worse outcome.[11,23]

The health economic analysis also identified that UDS is probably not cost-effective. This means that the service delivery offering of UDS can be reviewed, refocusing service priorities towards other indications.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Chapter 10 Research recommendations

UPSTREAM identified those men undergoing surgery to relieve BOO, but did not evaluate the quality of surgery received. It is possible that interventions did not fully relieve BOO in all cases. Further research is needed to identify the extent to which the quality of intervention influenced conclusions about the usefulness of preceding diagnostic tests.

Comorbidities are commonly present in men within the age range typically experiencing LUTS. UPSTREAM identified that comorbidity is influential on outcomes of interventional treatment. However, the assessment of comorbidities was not undertaken in detail. Further research should focus on a detailed assessment of comorbidity in order to establish what conditions are most relevant. Research should also attempt to explain why they are important; for example, by identifying relevant mechanisms, such as detrusor under activity or occult neurological disease.

Storage LUTS are common in older men, often coexisting with voiding symptoms. In UPSTREAM, storage LUTS were evaluated using the incontinence subscore of the ICIQ-MLUTS PROM, and the presence of detrusor overactivity in those men undergoing UD tests. These did not identify the predictive relevance of storage LUTS. Nonetheless, surgeons typically regard storage LUTS as an adverse predictive factor. Further research is needed to ascertain the most appropriate way to evaluate storage LUTS in this context, and to identify whether they are predictive of the outcome of operations to relieve BOO.

Post-void residual is sometimes used by surgeons as a reason to recommend surgery to relieve BOO. Nonetheless, the presence of PVR shows intra-individual variability, and uncertainty about the reliability of measurement accuracy. Furthermore, the underlying pathophysiological processes may reflect influences such as both BOO and detrusor under activity. Further research should identify subgroups of people with PVR, and whether those men who have BOO without underactivity have a better chance of reduced PVR following surgery to relieve BOO.

Chapter 11 Conclusions

This 5-year follow-up study found minimal difference between the two groups in terms of urinary symptoms and surgery rates. This supports the original findings at 18 months that symptomatic outcome is non-inferior in the UD group, and that the key secondary outcome of reduced surgery rates was not seen. The health economic analysis identified that the inclusion of UDS in the assessment pathway of men considering surgery for LUTS is probably not cost-effective.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Additional information

CRediT contribution statement

Madeleine Clout (https://orcid.org/0000-0001-7645-1199): Project administration, Writing – reviewing and editing.

Amanda L Lewis (https://orcid.org/0000-0003-0488-5347): Project administration, Writing – original draft.

Madeleine Cochrane (https://orcid.org/0000-0003-1856-3293): Data curation, Formal analysis, Writing – reviewing and editing, Visualisation.

Grace J Young (https://orcid.org/0000-0002-5210-1183): Data curation, Formal analysis, Writing – reviewing and editing, Visualisation.

Paul Abrams (https://orcid.org/0000-0003-2776-2200): Funding acquisition, Writing – reviewing and editing.

Peter S Blair (https://orcid.org/0000-0002-7832-8087): Funding acquisition, Writing – reviewing and editing.

Christopher Chapple (https://orcid.org/0000-0002-2960-9931): Funding acquisition, Writing – reviewing and editing.

Gordon T Taylor (https://orcid.org/0000-0002-8029-2503): Funding acquisition, Writing – reviewing and editing.

Sian Noble (https://orcid.org/0000-0002-8011-0722): Data curation, Formal analysis, Writing – reviewing and editing.

Tom Steuart-Feilding (https://orcid.org/0000-0002-4707-7828): Project administration.

Jodi Taylor (https://orcid.org/0000-0001-7171-8923): Conceptualisation, Funding acquisition, Writing – reviewing and editing.

J Athene Lane (https://orcid.org/0000-0002-7578-4925): Conceptualisation, Funding acquisition, Methodology, Writing – reviewing and editing.

Marcus J Drake (https://orcid.org/0000-0002-6230-2552): Conceptualisation, Funding acquisition, Methodology, Writing – reviewing and editing.

Acknowledgements

The authors wish to thank all the men who participated in UPSTREAM (Phase I and Phase II). We would also like to thank all 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 – a list of participating centres is reported elsewhere. The study would not have been possible without the involvement of patients and the centres; we are eternally grateful.

We also thank the members of the independent TSC [Professor Mark Emberton, Consultant Urological Surgeon (chairperson); Mr Ian Eardley, Consultant Urological Surgeon; Mr Andrew Elders, Senior Statistician; and Dr Malcolm Pick, patient and public involvement representative] for their continued counsel and support.

We are also grateful to Amanda Lewis for her Trial Management support from 2014 through to June 2022 and Tom Steuart-Feilding for his administrative support during UPSTREAM Phase II.

We wish to acknowledge Professor Rob Pickard, a consultant urological surgeon and co-applicant, who sadly passed away in July 2018. Professor Pickard was a major contributor in developing the original funding application and contributed valuable expertise to the study design and oversight of UPSTREAM Phase I.

Finally, we thank the NIHR and the HTA programme for funding UPSTREAM (reference number 12/140/01), and NBT for study sponsorship [reference number 4560 (Phase II)].

This study was designed and delivered in collaboration with the Bristol Trials Centre, a UK Clinical Research Collaboration registered clinical trials unit that was in receipt of NIHR Clinical Trials Units support funding.

Study data were collected and managed using REDCap, hosted at the University of Bristol.

Patient data statement

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it is important that there are safeguards to make sure that they are stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation

Data-sharing statement

Requests for access to data should be addressed to the corresponding author or to the data custodian (if known).

Ethics statement

The Proportionate Review Sub-Committee of the South Central – Berkshire Research Ethics Committee reviewed and approved UPSTREAM study (Phase II) on 1 November 2019 (REC reference 19/SC/0578).

Information governance statement

North Bristol NHS Trust and the University of Bristol are committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679.

Under the Data Protection legislation, North Bristol NHS Trust and the University of Bristol are joint Data Controllers for the UPSTREAM study. You can find out more about how these organisations handle personal data, including how to exercise your individual rights and the contact details for the Data Protection Officers here:

www.nbt.nhs.uk/about-us/information-governance/privacy-policy-data-protection

www.bristol.ac.uk/secretary/data-protection/gdpr/data-protection-officer/

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/SLPT4675.

Primary conflicts of interest: Paul Abrams reports personal fees from Astellas, Coloplast, Pierre Fabre, Tillotts, Sun Pharma, Cipla and Sandoz, outside of the submitted work, and is the Chair of the International Consultation on Incontinence – Research Society. J Athene Lane reports membership of the NIHR CTU Standing Advisory Committee during the submitted work. Marcus J Drake reports personal fees from Astellas and Pfizer, outside of the submitted work, and is a Trustee of the International Continence Society.

Publications

Publications (peer-reviewed)

1. **Background:** Drake MJ, Lewis AL, Lane JA. Urodynamic testing for men with voiding symptoms considering interventional therapy: the merits of a properly constructed randomised trial. *Eur Urol* 2016;**69**:759–60. https://doi.org/10.1016/j.eururo.2016.01.035

2. **Protocol:** 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. https://doi.org/10.1186/s13063-015-1087-1

3. **Statistical analysis plan:** 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**:455. https://doi.org/10.1186/s13063-017-2206-y

4. **Interview findings (1):** Selman LE, Ochieng CA, Lewis AL, Drake MJ, Horwood J. Recommendations for conducting invasive urodynamics for men with lower urinary tract symptoms: qualitative interview findings from a large randomized controlled trial (UPSTREAM). *Neurourol Urodyn* 2019;**38**:320–29.

5. **Baseline outcomes:** 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* 2019;**5**:340–50.

6. **Quality control:** Aiello M, Jelski J, Lewis A, Worthington J, McDonald C, Abrams P, *et al.* Quality control of uroflowmetry and urodynamic data from two large multicentre studies of male lower urinary tract symptoms. *Neurourol Urodyn* 2020;**39**:1170–7. https://doi.org/10.1002/nau.24337

7. **IPSS and ICIQ grading severity.** Ito I, Young GJ, Lewis AL, Blair P, Cotterill N, Lane JA, *et al.* Grading severity and bother using the IPSS and ICIQ-MLUTS scores in men seeking lower urinary tract symptoms therapy. *J Urol* 2020;**204**:1003–11. https://doi.org/10.1097/JU.0000000000001149

8. **Main results:** Drake MJ, Lewis AL, Young GJ, Abrams P, Blair P, Chapple C, *et al.* Diagnostic assessment of lower urinary tracts symptoms in men considering prostate surgery; a non-inferiority randomised controlled trial of urodynamics in 26 hospitals. *Eur Urol* 2020;**78**:701–10. https://doi.org/10.1016/j.eururo.2020.06.004

9. **HTA report.** Lewis AL, Young GJ, Selman LE, Rice C, Clement C, Ochieng CA, *et al.* Urodynamics tests for the diagnosis and management of bladder outlet obstruction in men: the UPSTREAM non-inferiority RCT. *Health Technol Assess* 2020;**24**:1–122.

10. **Interview findings (2):** Selman LE, Clement C, Ochieng CA, Lewis AL, Chapple C, Abrams P, *et al.* Treatment decision-making among men with lower urinary tract symptoms: a qualitative study of men's experiences with recommendations for patient-centred practice. *Neuro Urodyn* 2020;**40**:201–10.

11. **Exploratory analysis (1).** Ito H, Abrams P, Lewis AL, Young GJ, Blair PS, Cotterill N, *et al.* Use of the International Consultation on Incontinence Questionnaires bladder diary (ICIQ-BD) in men seeking therapy for lower urinary tract symptoms (LUTS). *Eur Urol Focus* 2022;**8**:66–74.

12. **Exploratory analysis (2).** Young GJ, Metcalfe C, Lane A, Lewis AL, Abrams P, Blair PS, *et al.* Prostate surgery for men with lower urinary tract symptoms: do we need urodynamics to find the right candidates? Exploratory findings from the UPSTREAM trial. *Eur Urol Focus* 2021;**S2405-4569**:00310-2. https://doi.org/10.1016/j.euf.2021.11.010

13. Clout M, Lewis AL, Cochrane M, Young GJ, Abrams P, Blair PS, *et al*. Five-year Follow-up of the Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods—A Noninferiority Randomised Controlled Trial of Urodynamics [published online ahead of print March 14 2025]. *Eur Urol Focus* 2025. https://doi.org/10.1016/j.euf.2025.02.004

Additional publications to follow (from Phase I)

13. **Health economics (1):** Main cost-effectiveness analysis *(title TBC)*

Target journal: *European Urology*

◦ To include Phase II, thus not expected until 2024+

14. **Health economics (2):** Methodology *(title TBC)*

Target journal: *Value in Health*

◦ To include Phase II, thus not expected until 2024+

Conference abstracts/presentations

European Association of Urology meeting 2023 state-of-the-art lecture and American Urological Association 2023 plenary lecture.

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

References

1. Lewis AL, Young GJ, Selman LE, Rice C, Clement C, Ochieng CA, *et al*. Urodynamics tests for the diagnosis and management of bladder outlet obstruction in men: the UPSTREAM non-inferiority RCT. *Health Technol Assess* 2020;**24**:1–122.
2. 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.
3. Jacobsen SJ, Girman CJ, Guess HA, Rhodes T, Oesterling JE, Lieber MM. Natural history of prostatism: longitudinal changes in voiding symptoms in community dwelling men. *J Urol* 1996;**155**:595–600.
4. 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**:455.
5. 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**:340–50.
6. Selman LE, Ochieng CA, Lewis AL, Drake MJ, Horwood J. Recommendations for conducting invasive urodynamics for men with lower urinary tract symptoms: qualitative interview findings from a large randomized controlled trial (UPSTREAM). *Neurourol Urodyn* 2019;**38**:320–9.
7. Gravas S, Cornu JN, Drake MJ, Gaccie C, Gratzke C, Herrmann S, *et al*. EAU guidelines on the assessment of non-neurogenic male lower urinary tract symptoms including benign prostatic obstruction. *Eur Urol* 2018; **67**: 1110.
8. Clout M, Lewis AL, Cochrane M, Young GJ, Abrams P, Blair PS, *et al*. Five-year Follow-up of the Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods—A Noninferiority Randomised Controlled Trial of Urodynamics [published online ahead of print March 14 2025]. *Eur Urol Focus 2025*. https://doi.org/10.1016/j.euf.2025.02.004
9. Barry MJ, Fowler FJ Jr, O'Leary MP, Bruskewitz RC, Holtgrewe HL, Mebust WK, Cockett AT. The American Urological Association symptom index for benign prostatic hyperplasia. The measurement committee of the American Urological Association. *J Urol* 1992;**148**:1549–57; discussion 1564.
10. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010;**8**:18.
11. Young GJ, Metcalfe C, Lane JA, Lewis AL, Abrams P, Blair PS, *et al*. Prostate surgery for men with lower urinary tract symptoms: do we need urodynamics to find the right candidates? Exploratory findings from the UPSTREAM trial. *Eur Urol Focus* 2022;**8**:1331–9.
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**:701–10.
13. Cochrane M, Noble S, Drake MJ. The UPSTREAM Study Phase II: Health Economic Analysis Plan (HEAP). 2023.
14. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013*. 2013. URL: www.nice.org.uk/process/pmg9/chapter/the-reference-case (accessed 1 November 2024).
15. Thorn JC, Davies CF, Brookes ST, Noble SM, Dritsaki M, Gray E, *et al*. Content of health economics analysis plans (HEAPs) for trial-based economic evaluations: expert Delphi consensus survey. *Value Health* 2021;**24**:539–47.

16. 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**:1099–109.

17. NHS England Digital. *HRG4+ 2020/21 National Costs Grouper*. 2021. URL: https://digital.nhs.uk/services/national-casemix-office/downloads-groupers-and-tools/hrg4-2020-21-national-costs-grouper (accessed 1 November 2024).

18. NHS England. *National Cost Collection: National Schedule of NHS Costs 2020/21*. 2022. URL: www.england.nhs.uk/publication/2020-21-national-cost-collection-data-publication/ (accessed 2 June 2025).

19. NHS England and NHS Improvement. *2020/21 National Tariff Payment System*. London: NHS England and NHS Improvement; 2020.

20. National Institute for Health and Care Excellence. *Developing NICE Guidelines: The Manual*. 2014 [updated 21 November 2023]. URL: www.nice.org.uk/process/pmg20/chapter/incorporating-economic-evaluation (accessed 1 November 2024).

21. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, *et al*. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;**15**:708–15.

22. National Institute for Health and Care Research. *Randomised Controlled Trial Participants: Diversity Data Report*. 2022. URL: www.nihr.ac.uk/documents/randomised-controlled-trial-participants-diversity-data-report/31969 (accessed 1 November 2024).

23. Ito H, Sakamaki K, Young GJ, Blair PS, Hashim H, Lane JA, *et al*. Predicting prostate surgery outcomes from standard clinical assessments of lower urinary tract symptoms to derive prognostic symptom and flowmetry criteria. *Eur Urol Focus* 2023;**10**:197–204.

24. Aiello M, Jelski J, Lewis A, Worthington J, McDonald C, Abrams P, *et al*. Quality control of uroflowmetry and urodynamic data from two large multicenter studies of male lower urinary tract symptoms. *Neurourol Urodyn* 2020;**39**:1170–7.

25. 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**:1003–11.

26. Selman LE, Clement C, Ochieng CA, Lewis AL, Chapple C, Abrams P, *et al*. Treatment decision-making among men with lower urinary tract symptoms: a qualitative study of men's experiences with recommendations for patient-centred practice. *Neurourol Urodyn* 2021;**40**:201–10.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 1 Baseline characteristics

TABLE 23 Baseline characteristics for those who did/didn't complete the 5-year questionnaire

	n[a]	Not included in the 5-year questionnaire analysis	*n*[a]	Included in the 5-year questionnaire analysis
Age				
Age at randomisation (years)	398	*67.99 (10.52)*	415	*67.33 (7.76)*
Ethnicity				
White	389	357 (92%)	404	372 (92%)
Other ethnic group	389	32 (8%)	404	32 (8%)
IMD score (based on postcodes)				
Median IMD score 2015 (IQR)[b]	390	*14* (8–24)	404	14 (8–22)
Quintile 1 *(most deprived)*		61 (16%)		43 (11%)
Quintile 2		60 (15%)		64 (16%)
Quintile 3	390	84 (22%)	404	99 (25%)
Quintile 4		98 (25%)		94 (23%)
Quintile 5 *(least deprived)*		87 (22%)		104 (26%)
Clinical baseline characteristics				
Comorbidities at baseline	394	278 (71%)	409	263 (64%)
Digital rectal examination findings[c]				
No abnormality	378	111 (29%)	401	117 (29%)
Benign enlargement	378	285 (75%)	401	308 (77%)
Suspected prostate cancer	378	10 (3%)	401	13 (3%)
Other	378	25 (7%)	401	17 (4%)
Uroflowmetry				
Maximum flow rate – Q_{max} (ml/second)	369	*12.61 (7.54)*	404	11.82 (6.54)
Median PVR in ml (IQR)	371	100 (38–191)	403	96 (45–173)
Median voided volume in ml (IQR)	376	204 (132.5–292.5)	405	228 (156–333)
Additional tests				
PSA test	383	58 (15%)	413	56 (14%)
Cystoscopy	383	42 (11%)	413	27 (7%)
Urinalysis	383	69 (18%)	413	49 (12%)
Kidney ultrasound	383	15 (4%)	413	10 (2%)
Cytology	383	1 (< 1%)	413	3 (1%)
Prostate volume measurement	383	9 (2%)	413	13 (3%)
Urea and electrolytes	383	22 (6%)	413	13 (3%)

TABLE 23 Baseline characteristics for those who did/didn't complete the 5-year questionnaire (*continued*)

	n[a]	Not included in the 5-year questionnaire analysis	*n*[a]	Included in the 5-year questionnaire analysis
IPSS: symptom severity at baseline				
Incomplete emptying	377	2.85 (1.74)	413	2.66 (1.69)
Frequency	377	3.52 (1.39)	413	3.40 (1.27)
Intermittency	377	2.69 (1.71)	413	2.55 (1.60)
Urgency	377	2.86 (1.71)	411	2.55 (1.62)
Weak stream	376	3.12 (1.61)	412	3.21 (1.56)
Straining	373	1.62 (1.70)	412	1.60 (1.53)
Nocturia	377	2.74 (1.34)	412	2.58 (1.26)
Total IPSS score	366	19.37 (7.05)	408	18.55 (6.99)
IPSS QoL score	378	4.29 (1.33)	412	3.99 (1.29)
ICIQ MLUTS				
Voiding score	360	8.95 (4.32)	404	9.21 (4.11)
Incontinence score	364	5.28 (3.55)	400	4.93 (3.08)
Daytime frequency (> 8 times)	367	170 (46%)	405	159 (39%)
Nocturia (> 1 times per night)	368	289 (79%)	404	312 (77%)
ICIQ MLUTS – sexual matters				
Erections (reduced or none)	350	262 (75%)	401	290 (72%)
Ejaculation (reduced or none)	348	279 (80%)	394	316 (80%)
Painful ejaculation (yes)	324	55 (17%)	378	72 (19%)
Urinary symptoms affected sex life?	344	229 (67%)	392	263 (67%)

a The number with available baseline data.
b Higher scores mean higher levels of deprivation (http://geoconvert.mimas.ac.uk).
c These were not treated as mutually exclusive and centre staff were asked to tick all that applied, the denominator is the number of men who had a DRE date.

Note
Data are mean (SD) or *n* (%), unless otherwise stated.

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 2 Phase I mortality

TABLE 24 Mortality, according to ONS death records, for Phase I (first 18 months)

	UDS[a] (*n* = 415)	RC[a] (*n* = 386)
Phase I causes of death		
Chronic ischaemic heart disease	1	2
Intracerebral haemorrhage	1	0
Malignant neoplasm of bronchus and lung	1	0
Malignant neoplasm of kidney, except renal pelvis	1	0
Malignant neoplasm of oesophagus	1	0
Malignant neoplasm of other connective and soft tissue	1	0
Malignant neoplasm of pancreas	1	1
Other interstitial pulmonary diseases	1	0
Pneumonitis due to solids and liquids	1	0
Subarachnoid haemorrhage	1	0
Viral pneumonia, not elsewhere classified	1	0
Total	*11*	*3*

a There were 802 participants whose details were requested from NHS England. However, we were informed that one participant (in the UD group) could not be traced.

Note
Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

Appendix 3 Phase II mortality

TABLE 25 Mortality, according to ONS death records, for Phase II (18 months to 5 years)

	UDS[a] (*n* = 415)	RC[a] (*n* = 386)
Phase II causes of death		
Acute and subacute endocarditis	0	1
Acute myocardial infarction	2	1
Alzheimer disease	0	1
Cerebral infarction	1	0
Chronic ischaemic heart disease	1	2
Emergency use of U07 (COVID-19)	2	0
Fall on and from ladder	1	0
Influenza due to identified seasonal influenza virus	1	1
Intracerebral haemorrhage	1	0
Lymphoid leukaemia	1	0
Malignant neoplasm of bronchus and lung	1	2
Malignant neoplasm of kidney, expect renal pelvis	0	1
Malignant neoplasm of oesophagus	0	2
Malignant neoplasm of other and ill-defined digestive organs	0	1
Malignant neoplasm of pancreas	1	1
Malignant neoplasm of prostate	0	1
Malignant neoplasm of rectosigmoid junction	1	0
Non-follicular lymphoma	0	1
Other degenerative diseases of nervous system, not elsewhere classified	0	1
Other disorders of bladder	0	1
Other peripheral vascular diseases	0	1
Paralytic ileus and intestinal obstruction without hernia	0	1
Pneumonia, organism unspecified	2	0
Spinal muscular atrophy and related syndromes	1	1
Stroke, not specified as haemorrhage or infarction	0	2
Unspecified fall	0	1
Vascular dementia	1	0
Total	*17*	*23*

a There were 802 participants whose details were requested from NHS England. However, we were informed that one participant (in the UD group) could not be traced.

Note

Reproduced with permission from Clout *et al.*[8] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) licence, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The table includes minor additions and formatting changes to the original text.

Copyright © 2025 Clout *et al.* This work was produced by Clout *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Appendix 4 Multiple imputation for economic evaluation

****

//Registering data set to be imputed//

****

mi set wide

//Cost variables to be imputed- registering the partially observed variables//

****Done in order of least missing***Missingness for costs and length of life is the same (*n* = 19 no cost and length of life data)

mi register imputed lifeyearsgained_from0m6m lifeyearsgained_from6m12m lifeyearsgained_from12m18m lifeyearsgained_from18m24m lifeyearsgained_from24mYr3 lifeyearsgained_fromYr3Yr4 lifeyearsgained_fromYr4Yr5 all_costs_0m6m_apc all_costs_6m12m_apc all_costs_12m18m_apc all_costs_18m24m_apc all_costs_Yr3_apc all_costs_Yr4_apc all_costs_Yr5_apc all_costs_0m6m_op all_costs_6m12m_op all_costs_12m18m_op all_costs_18m24m_op all_costs_Yr3_op all_costs_Yr4_op all_costs_Yr5_op

//EQ5D score at each time point (if we imputed QALYs we would have loss of information/inefficient) variables to be imputed- register the partially observed variables

****Done in order of least missing*** 18 months is registered first as it had less missingness than 6 months, 12 months and 5 years

mi register imputed EQ5D_index_18m EQ5D_index_6m EQ5D_index_12m EQ5D_index_5Yrs

//Fully observed variables as regular variables that do not require imputation (baseline covariates adjusted for in main analysis are centre, baseline IPSS score, baseline EQ-5D, also includes arm)

****Done in order of least missing***Though IPSS and EQ5D baseline scores have simple mean imputation.

mi register regular arm centre b_IPSS_score EQ5D_index_baseline

mi describe

****

//Running the multiple imputation

****

//Runs multiple imputation with chained equations with predictive mean matching (pmm) over 60 (add) imputations by treatment group, setting seed at 101 (rseed)

//i prefix for fully observed categorical data (e.g. Centre)

//Note, before the equals sign are the variables with missing data

```
mi impute chained (pmm, knn(5)) lifeyearsgained_from0m6m lifeyearsgained_from6m12m lifeyearsgained_
from12m18m lifeyearsgained_from18m24m lifeyearsgained_from24mYr3 lifeyearsgained_fromYr3Yr4
lifeyearsgained_fromYr4Yr5 all_costs_0m6m_apc all_costs_6m12m_apc all_costs_12m18m_apc all_costs_18m24m_apc
all_costs_Yr3_apc all_costs_Yr4_apc all_costs_Yr5_apc all_costs_0m6m_op all_costs_6m12m_op all_costs_12m18m_op
all_costs_18m24m_op all_costs_Yr3_op all_costs_Yr4_op all_costs_Yr5_op EQ5D_index_18m EQ5D_index_6m EQ5D_
index_12m EQ5D_index_5Yrs = i.centre b_IPSS_score EQ5D_index_baseline, add(60) by(arm) rseed(101)

**************************

//Calculate discounted QALYs at each year

**************************

//NB we could equally well apply the discounting before imputing but when I did this for this data, the model would
not run

*life-years gained from the two time points is already calculated as, for example 0.5 years, 1 year, < 0.5/1 year, or 0 if
died earlier on during the trial.

*EQ-5D has zero for 6 months, 18 months and 5 years if they died before that period.

//Calculate discounted QALYs at each year- 3.5%

mi passive: gen QALY_y1 =  (((EQ5D_index_baseline + EQ5D_index_6m)/2)*lifeyearsgained_from0m6m) + (((EQ5D_
index_6m + EQ5D_index_12m)/2)*lifeyearsgained_from6m12m)

mi passive: gen QALY_y2_D= (((EQ5D_index_12m + EQ5D_index_18m)/2)*lifeyearsgained_from12m18m) + (((EQ5D_
index_18m + EQ5D_index_5Yrs)/2)*lifeyearsgained_from18m24m)*((1.035)^(-1))

mi passive: gen QALY_y3_D = (((EQ5D_index_18m + EQ5D_index_5Yrs)/2)*lifeyearsgained_from24mYr3)*((1.035)^(-2))

mi passive: gen QALY_y4_D = (((EQ5D_index_18m + EQ5D_index_5Yrs)/2)*lifeyearsgained_fromYr3Yr4)*((1.035)^(-3))

mi passive: gen QALY_y5_D = (((EQ5D_index_18m + EQ5D_index_5Yrs)/2)*lifeyearsgained_fromYr4Yr5)*((1.035)^(-4))

**************************

//Calculate discounted QALYs at each year- 3.5%

**************************

//NB we could equally well apply the discounting before imputing but when I did this for this data, the model would
not run

//Calculate discounted costs at each year

mi passive: gen costs_y1 = (all_costs_0m6m_apc + all_costs_6m12m_apc + all_costs_0m6m_op + all_costs_6m12m_op)

mi passive: gen costs_y2_D = (all_costs_12m18m_apc + all_costs_18m24m_apc + all_costs_12m18m_op + all_
costs_18m24m_op)*((1.035)^(-1))

mi passive: gen costs_y3_D = (all_costs_Yr3_apc + all_costs_Yr3_op)*((1.035)^(-2))
```

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

mi passive: gen costs_y4_D = (all_costs_Yr4_apc + all_costs_Yr4_op)*((1.035)^(-3))

mi passive: gen costs_y5_D = (all_costs_Yr5_apc + all_costs_Yr5_op)*((1.035)^(-4))

*ONLY APC

mi passive: gen costs_apc_y1 = (all_costs_0m6m_apc + all_costs_6m12m_apc)

mi passive: gen costs_apc_y2_D = (all_costs_12m18m_apc + all_costs_18m24m_apc)*((1.035)^(-1))

mi passive: gen costs_apc_y3_D = (all_costs_Yr3_apc)*((1.035)^(-2))

mi passive: gen costs_apc_y4_D = (all_costs_Yr4_apc)*((1.035)^(-3))

mi passive: gen costs_apc_y5_D = (all_costs_Yr5_apc)*((1.035)^(-4))

mi passive: gen total_apc_costs_D = costs_apc_y1 + costs_apc_y2_D + costs_apc_y3_D + costs_apc_y4_D + costs_apc_y5_D//

*ONLY OP

mi passive: gen costs_op_y1 = (all_costs_0m6m_op + all_costs_6m12m_op)

mi passive: gen costs_op_y2_D = (all_costs_12m18m_op + all_costs_18m24m_op)*((1.035)^(-1))

mi passive: gen costs_op_y3_D = (all_costs_Yr3_op)*((1.035)^(-2))

mi passive: gen costs_op_y4_D = (all_costs_Yr4_op)*((1.035)^(-3))

mi passive: gen costs_op_y5_D = (all_costs_Yr5_op)*((1.035)^(-4))

mi passive: gen total_op_costs_D = costs_op_y1 + costs_op_y2_D + costs_op_y3_D + costs_op_y4_D + costs_op_y5_D// create

****

//Create variable for: Discounted Total QALYs and Total Costs

****

mi passive: gen total_QALYs = QALY_y1 + QALY_y2_D + QALY_y3_D + QALY_y4_D + QALY_y5_D//create variable for total QALYs

mi passive: gen total_costs = costs_y1 + costs_y2_D + costs_y3_D + costs_y4_D + costs_y5_D// create variable for total costs

replace arm= 0 if arm== 2

destring id, replace

*Discounted

```
///Total APC

mi estimate: regress total_apc_costs_D i.arm i.centre b_IPSS_score

regress total_apc_costs_D i.arm i.centre b_IPSS_score

///Total OP

mi estimate: regress total_op_costs_D i.arm i.centre b_IPSS_score

regress total_op_costs_D i.arm i.centre b_IPSS_score

///Year 1

mi estimate: regress costs_apc_y1 i.arm i.centre b_IPSS_score

mi estimate: regress costs_op_y1 i.arm i.centre b_IPSS_score

mi estimate: regress QALY_y1 i.arm i.centre b_IPSS_score EQ5D_index_baseline

///Year 2

mi estimate: regress costs_apc_y2_D i.arm i.centre b_IPSS_score

mi estimate: regress costs_op_y2_D i.arm i.centre b_IPSS_score

mi estimate: regress QALY_y2_D i.arm i.centre b_IPSS_score EQ5D_index_baseline

///Year 3

mi estimate: regress costs_apc_y3_D i.arm i.centre b_IPSS_score

mi estimate: regress costs_op_y3_D i.arm i.centre b_IPSS_score

mi estimate: regress QALY_y3_D i.arm i.centre b_IPSS_score EQ5D_index_baseline

///Year 4

mi estimate: regress costs_apc_y4_D i.arm i.centre b_IPSS_score

mi estimate: regress costs_op_y4_D i.arm i.centre b_IPSS_score

mi estimate: regress QALY_y4_D i.arm i.centre b_IPSS_score EQ5D_index_baseline

///Year 5

mi estimate: regress costs_apc_y5_D i.arm i.centre b_IPSS_score

mi estimate: regress costs_op_y5_D i.arm i.centre b_IPSS_score

mi estimate: regress QALY_y5_D i.arm i.centre b_IPSS_score EQ5D_index_baseline
```

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

```
**************************

*Seemingly Unrelated Regression: Analysis of multiple imputed data sets (mi impute chained)

**************************

//Regress using seemingly unrelated regression (SUR)//

*cmdok allow estimation when estimation command is not one of the officially supported estimation commands

mi estimate, cmdok: sureg (total_costs i.arm i.centre b_IPSS_score) (total_QALYs i.arm i.centre b_IPSS_score EQ5D_
index_baseline), corr

mimrgns arm, predict(equation(total_costs))

mimrgns arm, predict(equation(total_QALYs))

**************************

*PER YEAR: Seemingly Unrelated Regression: Analysis of multiple imputed data sets (mi impute chained)

**************************

//Year 1

mi estimate, cmdok: sureg (costs_y1 i.arm i.centre b_IPSS_score) (QALY_y1 i.arm i.centre b_IPSS_score EQ5D_index_
baseline), corr

mimrgns arm, predict(equation(costs_y1))

mimrgns arm, predict(equation(QALY_y1))

//Year 2

mi estimate, cmdok: sureg (costs_y2_D i.arm i.centre b_IPSS_score) (QALY_y2_D i.arm i.centre b_IPSS_score EQ5D_
index_baseline), corr

mimrgns arm, predict(equation(costs_y2_D))

mimrgns arm, predict(equation(QALY_y2_D))

//Year 3

mi estimate, cmdok: sureg (costs_y3_D i.arm i.centre b_IPSS_score) (QALY_y3_D i.arm i.centre b_IPSS_score EQ5D_
index_baseline), corr

mimrgns arm, predict(equation(costs_y3_D))

mimrgns arm, predict(equation(QALY_y3_D))
```

```
//Year 4

mi estimate, cmdok: sureg (costs_y4_D i.arm i.centre b_IPSS_score) (QALY_y4_D i.arm i.centre b_IPSS_score EQ5D_
index_baseline), corr

mimrgns arm, predict(equation(costs_y4_D))

mimrgns arm, predict(equation(QALY_y4_D))

//Year 5

mi estimate, cmdok: sureg (costs_y5_D i.arm i.centre b_IPSS_score) (QALY_y5_D i.arm i.centre b_IPSS_score EQ5D_
index_baseline), corr

mimrgns arm, predict(equation(costs_y5_D))

mimrgns arm, predict(equation(QALY_y5_D))

**********

//Probability of cost-effectiveness using coefficients from SUR

**********

//this is creating the INMB parametrically from the estimates of the sureg//

mi estimate, cmdok: sureg (total_costs i.arm i.centre b_IPSS_score) (total_QALYs i.arm i.centre b_IPSS_score EQ5D_
index_baseline), corr
```

Copyright © 2025 Clout *et al*. This work was produced by Clout *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: https://creativecommons.org/licenses/by/4.0/. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health and Care Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care

Published by the NIHR Journals Library