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

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

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

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

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

*Corresponding author

Health Technology Assessment

HTA/HTA TAR

> ## Criteria for inclusion in the *Health Technology Assessment* journal
> Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.
>
> Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

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

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

This report

Abstract

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

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

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

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

Background: Lower urinary tract symptoms (LUTS) in men may indicate bladder outlet obstruction (BOO) or weakness, known as detrusor underactivity (DU). Severe bothersome LUTS are a common indication for surgery. The diagnostic tests may include urodynamics (UDS) to confirm whether BOO or DU is the cause, potentially reducing the number of people receiving (inappropriate) surgery.

Objectives: The primary objective was to determine whether a care pathway including UDS is no worse for symptom outcome than one in which it is omitted, at 18 months after randomisation. Rates of surgery was the key secondary outcome.

Design: This was a pragmatic, multicentre, two-arm (unblinded) randomised controlled trial, incorporating a health economic analysis and qualitative research.

Setting: Urology departments of 26 NHS hospitals in England.

Participants: Men (aged ≥ 18 years) seeking further treatment, potentially including surgery, for bothersome LUTS. Exclusion criteria were as follows: unable to pass urine without a catheter, having a relevant neurological disease, currently undergoing treatment for prostate or bladder cancer, previously had prostate surgery, not medically fit for surgery and/or unwilling to be randomised.

Interventions: Men were randomised to a care pathway based on non-invasive routine tests (control) or routine care plus invasive UDS (intervention arm).

Main outcome measures: The primary outcome was International Prostate Symptom Score (IPSS) at 18 months after randomisation and the key secondary outcome was rates of surgery. Additional secondary outcomes included adverse events (AEs), quality of life, urinary and sexual symptoms, UDS satisfaction, maximum urinary flow rate and cost-effectiveness.

Results: A total of 820 men were randomised (UDS, 427; routine care, 393). Sixty-seven men withdrew before 18 months and 11 died (unrelated to trial procedures). UDS was non-inferior to routine care for IPSS 18 months after randomisation, with a confidence interval (CI) within the margin of 1 point (–0.33, 95% CI –1.47 to 0.80). A lower surgery rate in the UDS arm was not found (38% and 36% for UDS and routine care, respectively), with overall rates lower than expected. AEs were similar between the arms at 43–44%. There were more cases of acute urinary retention in the routine care arm. Patient-reported outcomes for LUTS improved in both arms and satisfaction with UDS was high in men who received it. UDS was more expensive than routine care. From a secondary care perspective, UDS cost an additional £216 over an 18-month time horizon. Quality-adjusted life-years (QALYs) were similar, with a QALY difference of 0.006 in favour of UDS over 18 months. It was established that UDS was acceptable to patients, and valued by both patients and clinicians for its perceived additional insight into the cause and probable best treatment of LUTS.

Limitations: The trial met its predefined recruitment target, but surgery rates were lower than anticipated.

Conclusions: Inclusion of UDS in the diagnostic tests results in a symptom outcome that is non-inferior to a routine care pathway, but does not affect surgical rates for treating BOO. Results do not support the routine use of UDS in men undergoing investigation of LUTS.

Future work: Focus should be placed on indications for selective utilisation of UDS in individual cases and long-term outcomes of diagnosis and therapy.

Trial registration: Current Controlled Trials ISRCTN56164274.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

AE	adverse event	ICIQ-MLUTSsex	International Consultation on Incontinence Questionnaire – Sexual Matters associated with Male Lower Urinary Tract Symptoms
BCI	Bladder Contractility Index		
BNI	bladder neck incision		
BOO	bladder outlet obstruction		
BOOI	Bladder Outlet Obstruction Index	ICIQ-UDS-S	International Consultation on Incontinence Questionnaire – urodynamics satisfaction
BPE	benign prostate enlargement		
BPO	benign prostatic obstruction	ICS	International Continence Society
BRTC	Bristol Randomised Trials Collaboration		
		IDUF	increased daytime urinary frequency
CACE	complier-average causal effect		
CEAC	cost-effectiveness acceptability curve	INMB	incremental net monetary benefit
		IPSS	International Prostate Symptom Score
CI	confidence interval		
CRF	case report form	ITT	intention to treat
DMC	Data Monitoring Committee	LUT	lower urinary tract
DRE	digital rectal examination	LUTS	lower urinary tract symptoms
DU	detrusor underactivity	NBT	North Bristol NHS Trust
EAU	European Association of Urology	NICE	National Institute for Health and Care Excellence
EQ-5D-5L	EuroQol-5 Dimensions, five-level version	NIHR	National Institute for Health Research
FCE	finished consultant episode	OAB	overactive bladder syndrome
GP	general practitioner	PROM	patient-reported outcome measure
HoLEP	holmium laser enucleation of the prostate	PSA	prostate-specific antigen
		PVR	post-void residual
HRG	Healthcare Resource Group	QALY	quality-adjusted life-year
HTA	Health Technology Assessment	Q_{max}	maximum urinary flow rate
ICER	incremental cost-effectiveness ratio	QoL	quality of life
		SAE	serious adverse event
ICIQ-MLUTS	International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms	SAP	statistical analysis plan
		SD	standard deviation

SUR	seemingly unrelated regression	UPSTREAM	Urodynamics for Prostate Surgery Trial: Randomised Evaluation of Assessment Methods
TMG	Trial Management Group		
TSC	Trial Steering Committee		
TURP	transurethral resection of the prostate	UTI	urinary tract infection
		VV	voided volume
UDS	urodynamics		

Plain English summary

After hospital referral, men with bothersome lower urinary tract symptoms (LUTS) are assessed with standard tests. These include measurement of urine flow rate, bladder diaries and questionnaires, including the International Prostate Symptom Score (IPSS). UPSTREAM (Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods) researched whether or not including an extra test, urodynamics (UDS), helps when considering treatment options. UDS is a more invasive test and measures pressure in the bladder to check whether or not the prostate is causing obstruction. It was presumed that, if there is no obstruction, surgery would not be offered, so that using UDS would reduce the number of prostate operations.

Each man participating (820 in total) was assessed with the standard tests. Around half of them had no extra tests (the 'routine care' arm of the trial); the rest had the UDS tests (the 'UDS' arm). Men then went on to have treatment, which they chose having discussed their test results with a urologist. IPSS and other symptom scores were examined for each man 18 months after joining the trial. At 18 months, surgery outcomes were known for 792 men and IPSS was known for 669 men.

We investigated if the two trial arms showed similar changes in the IPSS and if there were fewer operations done in the UDS arm. We identified similar reductions in the IPSS in both arms. However, UDS tests did not reduce the number of operations. Analysing all the costs, it was found that a pathway including UDS costs more than routine care.

Interviews were conducted that showed that men found UDS acceptable, and that the additional information helped both the men and their doctors consider which treatment would be most appropriate.

These results do not support the routine use of UDS in the assessment of every man considering prostate surgery for LUTS. Further exploration of the data may identify circumstances in which UDS could be helpful.

Scientific summary

Background

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

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

Objectives

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

The trial addressed several other key questions:

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

The qualitative component considered the following questions:

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

Methods

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

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

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

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

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

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

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

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

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

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

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

Primary analyses

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

Secondary analyses

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

Cost-effectiveness analysis

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

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

Qualitative evaluation

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

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

Results

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

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

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

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

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

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

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

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

Conclusions

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

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

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

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

Recommendations for research

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

Trial registration

This trial is registered as ISRCTN56164274.

Funding

Chapter 1 Introduction

Background

The lower urinary tract (LUT) serves to store urine, relying on the bladder to act as a reservoir. People need to pass urine (known as voiding); this involves the expulsion of urine through the bladder outlet, which is a tube called the urethra. When storing urine, the urethra is held shut by a muscle called the sphincter. When the person wishes to void, they do so by relaxing the sphincter and increasing their bladder pressure by contraction of the main bladder muscle, called the detrusor, so that urine is actively expelled. In men, the genital tract also uses the urethra, and an important sexual gland, the prostate, encircles the urethra between the bladder neck and the sphincter.

Lower urinary tract symptoms (LUTS) are classified according to their relationship to the two main LUT functions of urine storage and voiding. The main storage LUTS are increased daytime urinary frequency (IDUF), nocturia, urgency and incontinence. The main voiding symptoms are slow stream, intermittency, hesitancy, straining and terminal dribbling. Some symptoms occur shortly after voiding has concluded. These 'post-voiding LUTS' are post-micturition dribble and a feeling of incomplete emptying. LUTS are a common clinical feature of ageing, with a high proportion of men aged > 50 years reporting at least one LUTS[1] sufficient to impair quality of life (QoL), occupation and other activities.

Several mechanisms can give rise to LUTS:

- Storage LUTS can arise if the bladder becomes 'overactive', leading to urgency and IDUF, and sometimes nocturia.
- Storage LUTS can also arise if a patient develops an inflammation in the LUT, such as a urinary tract infection (UTI) or other abnormality.
- Nocturia may occur if there is excessive urine production from the kidneys overnight, a problem that can be caused by a patient's own habits or a range of medical problems.
- Voiding and post-voiding LUTS might result if the prostate enlarges to constrict the urethra. Benign prostate enlargement (BPE) with ageing may cause partial bladder outlet obstruction (BOO), a situation known as benign prostatic obstruction (BPO). BOO can also be caused by pathologies narrowing the outlet, such as urethral stricture or bladder neck contracture.
- Voiding and post-voiding LUTS can also be caused if the expulsion strength is impaired by weakening of the bladder, known as detrusor underactivity (DU).

Severe LUTS may require surgical treatment, which is one of the more common indications for surgery in the UK NHS, generally using transurethral resection of the prostate (TURP) or laser-based methods of enucleation or ablation. In 2016/17, > 18,000 TURPs were reported in England.[2] TURP is associated with morbidities such as blood loss, erectile dysfunction or incontinence, and a mortality rate of up to 0.25%.[3] Delayed-onset complications include urethral stricture and bladder neck contracture, and may affect almost 10% of men.[3]

Each patient can have one or more of these causes of LUTS present; understanding of the underlying mechanisms is needed to optimise their treatment. Tests used in clinical practice to understand the issues in each individual with male LUTS include the following:[4,5]

- Symptom scores are used to catalogue the LUTS present and their impact on the patient.
- A digital rectal examination (DRE) is done to palpate outward enlargement of the prostate; this gives an indication of BPE but does not establish that BPO is present, as obstruction is caused by inward enlargement, compressing the urethra.

- Urinalysis is undertaken to exclude infection.
- Uroflowmetry is a basic test of voiding, evaluating the maximum urinary flow rate (Q_{max}), voided volume (VV) and post-void residual (PVR). A threshold value for Q_{max} of 10 ml/second is commonly employed; for a well-conducted study, the specificity and positive predictive value for BOO may exceed 90%.[6] Other sources suggest that a Q_{max} of < 10 ml/second has a specificity of 70%, a positive predictive value of 70% and a sensitivity of 47% for BOO.[7]

An additional test that can contribute insight to the underlying mechanisms is urodynamics (UDS), a term that principally covers observations made during multichannel cystometry, including filling cystometry to assess storage function and pressure–flow studies of voiding function. UDS requires urethral catheterisation for bladder filling and pressure measurement. Anal catheterisation is also done for measurement of abdominal pressure from the rectum. During the test, a computer calculates 'detrusor pressure', derived by subtracting abdominal from bladder pressure, to demonstrate whether or not bladder contraction is occurring. This test is the only routine clinical test able to distinguish BOO from DU, based on observation of a slow Q_{max} with a high detrusor pressure[8] or with a low detrusor pressure,[9] respectively. The clinical benefit of UDS is to ensure that interventions aimed at reducing outlet obstruction are used only in men who actually have BPO and also to identify other risk factors. As men are unlikely to accept TURP if they do not have BPO, widespread use of UDS may reduce interventional therapy for BPO in the NHS.

Urodynamics is not routinely included in standard clinical male LUTS assessment pathways in the UK. The National Institute for Health and Care Excellence (NICE) clinical guideline for this condition suggests that UDS should be offered to men with LUTS having specialist assessment if they are considering surgery.[5] The European Association of Urology (EAU) suggested that pressure–flow studies may be used for patients who cannot void > 150 ml, have a PVR of > 300 ml and in men aged > 80 years, and should be performed in men aged < 50 years when considering surgery.[4] Both organisations alluded to the weak evidence base when developing recommendations, and this was also identified in a Cochrane systematic review[10] and other reviews.[11]

Rationale

For older men, a slow urinary stream is a common problem. It can be caused by partial BOO, meaning that the conduit for passing urine is restricted. The most common cause is prostate enlargement due to the nodular growth of the gland in response to the male sex hormones. As the prostate is located around the upper end of the urethra, any inward enlargement will impede flow at the top end of the urethra, referred to as BPO. If sufficient obstruction occurs, voiding LUTS will result. The enlargement of the gland also can occur outwards; this can be felt by physical examination, done by feeling the prostate (i.e. a DRE). Other possible causes of BOO include a stricture (constricting scar of the urethra) or failure of relaxation of the bladder neck.

Voiding LUTS of a similar nature can also be a result of weakness of the bladder when a man is attempting to pass urine. This effectively means that the bladder does not mount enough contraction strength to overcome the innate resistance of the normal male bladder outlet and any co-existing BOO will exacerbate the situation. The associated LUTS are referred to as underactive bladder, which is now formally defined by the International Continence Society (ICS): 'Underactive bladder is characterised by a slow urinary stream, hesitancy and straining to void, with or without a feeling of incomplete bladder emptying sometimes with storage symptoms'.[12]

Storage LUTS are also a common feature for older men. These can result from a few possible mechanisms. Two situations are particularly important:

1. Overactive bladder syndrome (OAB), which is a symptom syndrome also defined by the ICS: 'OAB is characterized by urinary urgency, with or without urgency urinary incontinence, usually with increased daytime frequency and nocturia, if there is no proven infection or other obvious pathology'.[13,14] By implication, this is a consequence of bladder dysfunction.

2. Nocturia, which is 'the complaint that the individual has to wake at night one or more times to void'.[14,15] This may reflect OAB, but can also be due to the production of large volumes of urine from the kidneys overnight, referred to as 'nocturnal polyuria'. Behavioural factors and medical problems can also give rise to nocturia, meaning that a symptom commonly referred to as a LUTS may actually reflect influences outside the LUT.[16,17]

Thus, LUTS result from a range of underlying causes, including BOO, bladder dysfunction and factors outside the LUT. For many individuals, LUTS are very bothersome and men attend for medical attention in the hope of alleviating the impact of their symptoms. Furthermore, these issues are highly prevalent.

The treatment of LUTS aims to deal with the underlying mechanism. Diagnostic pathways are set up to map the severity and impact of the LUTS, to exclude serious conditions that could give rise to apparently similar presentations and to suggest mechanism. Details of the NICE guidance on the management of LUTS in men[5] is provided in *Box 1*.

BOX 1 NICE guidance on management of LUTS in men

1.2 Specialist assessment

Specialist assessment refers to assessment carried out in any setting by a health-care professional with specific training in managing LUTS in men.

1.2.1 Offer men with LUTS having specialist assessment an assessment of their general medical history to identify possible causes of LUTS, and associated comorbidities. Review current medication, including herbal and over-the-counter medicines to identify drugs that may be contributing to the problem. [2010]

1.2.2 Offer men with LUTS having specialist assessment a physical examination guided by urological symptoms and other medical conditions, an examination of the abdomen and external genitalia, and a digital rectal examination (DRE). [2010]

1.2.3 At specialist assessment, ask men with LUTS to complete a urinary frequency volume chart. [2010]

1.2.4 At specialist assessment, offer men with LUTS information, advice and time to decide if they wish to have prostate specific antigen (PSA) testing if: their LUTS are suggestive of bladder outlet obstruction secondary to BPE or their prostate feels abnormal on DRE or they are concerned about prostate cancer. [2010]

1.2.5 Offer men with LUTS who are having specialist assessment a measurement of flow rate and post void residual volume. [2010]

1.2.6 Offer cystoscopy to men with LUTS having specialist assessment only when clinically indicated, for example if there is a history of any of the following:

- recurrent infection
- sterile pyuria
- haematuria
- profound symptoms
- pain. [2010]

BOX 1 NICE guidance on management of LUTS in men (*continued*)

1.2.7 Offer imaging of the upper urinary tract to men with LUTS having specialist assessment only when clinically indicated, for example if there is a history of any of the following:

- chronic retention
- haematuria
- recurrent infection
- sterile pyuria
- profound symptoms
- pain. [2010]

1.2.8 Consider offering multichannel cystometry to men with LUTS having specialist assessment if they are considering surgery. [2010]

1.2.9 Offer pad tests to men with LUTS having specialist assessment only if the degree of urinary incontinence needs to be measured. [2010]

The NICE guidance[5] also recommends that men considering any treatment for LUTS be offered an assessment of their baseline symptoms with a validated symptom score [e.g. the International Prostate Symptom Score (IPSS)], to allow assessment of subsequent symptom change. At initial assessment, the guidance[5] suggests that men with LUTS a urine dipstick test to detect blood, glucose, protein, leucocytes and nitrites. The EAU guidelines[4] include the same assessment components and, in addition, include the need to enquire about sexual function. Thus, two main diagnostic pathways relevant to practice in the UK include history and physical examination, including DRE; symptom score assessment; frequency–volume chart (also known as a bladder diary); flow rate testing with PVR measurement; and dipstick urinalysis. In both pathways, UDS testing is used selectively.

The potential contribution of UDS is principally in establishing if BOO is truly present. This is achieved by measuring the pressure the bladder generates when the flow rate is at its maximum. A high pressure generating only a slow flow is diagnostic of BOO and is quantified by the Bladder Outlet Obstruction Index (BOOI). A low pressure with slow flow rate indicates that reduced detrusor contraction strength is the cause, technically referred to as DU, and is quantified by the Bladder Contractility Index (BCI). Because of the nature of the way the UDS parameters are established, the test can identify if both mechanisms are present in one individual.

On the face of it, the confirmation of mechanism would make sense for guiding treatment choice. However, the prevalence of DU is lower than that of BPO; therefore, the chance of BPO as the explanation for voiding LUTS is considerably high. For men with a Q_{max} of < 10 ml/second, the chance that BOO is present is up to 90%, provided the test is undertaken and interpreted appropriately.[6] However, it becomes less reliable if the test is not done in adequate circumstances,[18] or if the flow rate is > 10 ml/second.

Thus, two broad approaches have emerged:

1. use of UDS included in the diagnostic pathway to confirm whether BOO and/or DU is the cause of voiding LUTS in an individual man
2. omission of UDS from the pathway and presuming that BPO is the problem for a man with voiding LUTS.

This has important implications for decision-making, as someone evaluated under the second approach could be considered for surgery to relieve presumed BPO when actually they do not have BOO. Consequently, they would potentially experience the intervention and possible complications, and not benefit from improved symptoms. On the other hand, men evaluated under the first approach will be expected to undergo the invasive diagnostic test of UDS, gaining little benefit if the finding simply backs up the suppositions derived by the other tests in the pathway. Thus, both pathways have potential advantages and disadvantages. Therefore, many urological centres use symptom assessment, physical examination, urinalysis, flow rate and PVR measurement, and bladder diary, hereafter called the 'routine care' diagnostic pathway. They choose to omit UDS, because of a lack of relevant research evidence. Indeed, a Cochrane review found only two studies, one of which had to be excluded from analysis.[10] Accordingly, NICE highlighted the importance of identifying the role of invasive UDS, including clarifying whether or not it could improve the outcome of surgery and whether or not it should be recommended in the future.[5]

When BPO is present or suspected, it can be treated by removing the obstruction. Relief of BOO can be attempted by various interventions. The following interventions were available during the course of the Urodynamics for Prostate Surgery Trial: Randomised Evaluation of Assessment Methods (UPSTREAM):

- medications to relax the urethra, known as alpha-blockers (e.g. tamsulosin, alfuzosin or doxazosin)
- medications to shrink the prostate enlargement, known as 5-alpha reductase inhibitors (e.g. finasteride or dutasteride)
- a medication whose mechanism of action in LUTS has not been fully established and the phosphodiesterase type 5 inhibitor tadalafil, which is also used in the treatment of erectile dysfunction
- endoscopic resection or incision of the intruding part of the prostate [TURP or bladder neck incision (BNI)]
- an endoscopic procedure to retract the prostate partly out of the way, known as UroLift® (NeoTract Inc., Pleasanton, CA, USA)
- laser enucleation of the nodular elements [e.g. holmium laser enucleation of the prostate (HoLEP)]
- laser vaporisation of the intruding part of the prostate (e.g. greenlight laser)
- prostate artery embolisation to shrink the prostate by reducing its blood supply.

UPSTREAM studied men with bothersome LUTS who were referred to urology departments when surgery was potentially being considered. Men were randomised to the specified intervention (UDS plus routine care) or comparator (routine care alone), and the trial was powered to ascertain significant difference in non-inferiority in symptoms at 18 months after randomisation. The key secondary outcome was surgery rates, as men in the UDS arm identified as not having BOO would not be recommended for surgery. A non-inferiority assessment was selected because men not receiving surgery would see no improvement or modest improvement in symptoms.

Transurethral resection of the prostate requires a median hospital stay of 2 days, and additional NHS costs can result from delayed discharge from hospital, readmissions and increased primary care utilisation. Significant risks may be associated with TURP: mortality is up to 0.25% and there is risk of blood loss, erectile dysfunction or incontinence. Late complications, notably urethral stricture, can affect almost 10% of men.[3] Lower surgery rates would potentially have the advantages of reduced adverse effects and resource use, so a health economic analysis was included. Qualitative interviewing was also employed to explore user acceptability and influences on decisions made by the participating men and the surgeons.

Aim and objectives

Using invasive UDS to categorise LUT function should improve patient selection for surgery compared with a pathway with no invasive UDS testing. Identifying men with LUTS who do not have BOO will reduce men's willingness to undergo surgery, which should reduce risk of surgical complications and substandard symptom outcomes.

Aim
The aim was to determine whether a care pathway including UDS is no worse for symptom outcome than one in which it is omitted, at 18 months after randomisation. The primary clinical outcome was the IPSS at 18 months after randomisation. Influence of UDS on rates of bladder outlet surgery was a main secondary outcome.

Objectives

- Does invasive UDS deliver similar or better symptomatic outcomes for LUTS measured by the IPSS at 18 months after randomisation?
- Does invasive UDS influence surgical decision-making, as reflected in differing surgery rates in the two diagnostic pathways?
- What is the cost-effectiveness of the two diagnostic pathways?
- What are the relative harms of UDS and the subsequent therapy?
- What subsequent NHS services are required (including repeat surgery or catheterisation for acute urinary retention) for men in each arm?
- What are the differential effects on QoL?

A qualitative component examined patients' and clinicians' views and experiences of UDS for male BOO and BOO surgery. The qualitative component considered the following questions:

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

Chapter 2 Methods

Trial design

This trial was a two-arm, multicentre, randomised controlled trial, randomising men seeking further treatment, which may include surgery, for their bothersome LUTS, to either a care pathway that includes UDS testing or a care pathway without it (i.e. routine care). UPSTREAM was designed as a non-inferiority trial to establish non-inferiority in symptom severity 18 months after randomisation, which was the primary outcome, measured using the IPSS. It derived from the assumption that UDS would decrease the need for surgery, as it would provide useful information regarding bladder function and better predict success of surgery. Surgery rates were measured (to establish superiority) as a key secondary outcome (*Figure 1*). Both the trial protocol[19] and statistical analysis plan (SAP)[20] are published and detail the trial background and intentions.

Ethics approval and research governance

The National Research Ethics Service Committee South Central – Oxford B reviewed and approved the trial on 10 July 2014 (reference number 14/SC/0237). Centre-specific assessments were completed by each of the participating NHS trusts (as listed in the protocol[19]) and all the necessary approvals were granted, including subsequent approvals for amendments (identified in *Table 1*). The trial was registered with the International Standard Randomised Controlled Trial Number (ISRCTN) registry (ISRCTN56164274) and sponsored by North Bristol NHS Trust (NBT) (reference number 3250).

The trial was conducted in accordance with the UK Research Governance Framework[21] and the recommendations of the Declaration of Helsinki.[22] All men who agreed to take part in the trial provided written informed consent and any adverse event (AE) classified (and confirmed) as serious, related and unexpected was reported to the sponsor and Research Ethics Committee within 15 days. All AEs were summarised and reviewed regularly by the sponsor and Trial Management Group (TMG), including the independent Data Monitoring Committee (DMC) and the Trial Steering Committee (TSC).

Amendments

Table 1 summarises the minor and substantial amendments to the trial after the original ethics approval was granted, all of which the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme supported. Amendments were made, following feedback from patients and centre staff, to improve understanding of, uptake of and involvement with the trial (e.g. clarifying the inclusion criteria). Furthermore, slower than anticipated recruitment rates at the beginning resulted in an extension (i.e. substantial amendment 3).
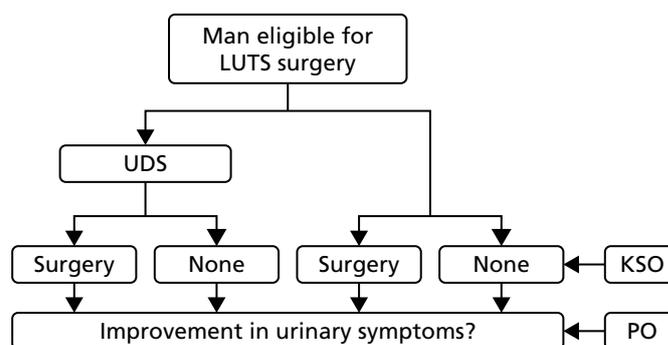
FIGURE 1 The UPSTREAM interventions and key outcomes. KSO, key secondary outcome; PO, primary outcome.

TABLE 1 Summary of amendments to the trial protocol since original ethics approval was granted

Amendment (date)	Brief description of amendment
Minor 1 (September 2014)	• Various documents updated to include contact details and improved formatting
Minor 2 (December 2015 and January 2016)	• Trial protocol: clarification of the 'data collection and transportation' process, subsequent AE reporting procedure and research nurse responsibilities. Trial timeline also updated, in connection with the point below • Notification of extended recruitment period (and subsequent 18-month follow-up) because of slower than anticipated rates across all participating centres. Note that the total trial duration remained unchanged from original REC application • Clarification regarding distribution of participant newsletters
Minor 3 (November 2016)	• Change of principal investigators at three centres: Freeman Hospital, Newcastle; Churchill Hospital, Oxford; and Royal Free Hospital, London
Minor 4 (April 2017)	• Change of principal investigator at one centre: Southport and Formby District General Hospital, Southport
Minor 5 (April 2017)	• Change of principal investigator at one centre: Churchill Hospital, Oxford
Substantial 1 (February 2015)	• Patient trial invitation letter revised, based on patient feedback. A brief overview of the trial was added to the letter to provide insight about why men were receiving the letter (invitation to the trial)
Substantial 2 (May 2015)	• Document revisions based on feedback, including the patient information sheet, consent form, protocol, questionnaires and letters/reminders. The changes were designed to improve the running of the trial by clarifying specific details that appeared to cause confusion (for both patients and staff, especially when identifying potential participants). Furthermore, the changes were made to improve a participant's experience, namely by reducing burden, providing more options for consent and methods of involvement, and clarifying processes that were unclear
Substantial 3 (September 2016)	• A 6-month contract variation was awarded by the funder (NIHR HTA) to extend recruitment for a further 6 months, resulting in a revised trial end date (i.e. 30 September 2018). Protocol amended, including changes to the way in which health economic and qualitative data were collected • Questionnaire reminder letter revised to include statement about the importance of return, regardless of patient treatment status. Prior to the change, for example, men who had completed treatment or not had surgery, appeared not to understand the importance of their responses, thinking 'it would not help'

AE, adverse event; HTA, Health Technology Assessment; NIHR, National Institute for Health Research; REC, Research Ethics Committee.

Patient and public involvement

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

During the funding application and development (design) of the trial, a user (patient) panel of eight volunteers from NBT's research and innovation department provided insight and feedback, and 15 men were consulted at flow rate/UDS clinics at two hospitals about the procedures that would, and could, occur during the trial. All these patient representatives regarded invasive UDS as acceptable, to gain information relevant to treatment decisions. Experiencing the test did not alter this attitude: it was 'less unpleasant than anticipated'. For a care pathway omitting invasive testing, all men understood how sufficient information for treatment could be obtained. One man said that 'free-flow rate testing is time-consuming' and recommended keeping the diagnostic pathway short. Thirteen of the 15 patient representatives would accept random allocation to a diagnostic pathway. One said, 'he did not want any tubes inserted without general anaesthetic'; another said he was 'convinced of the benefits of the test and a more accurate diagnosis'. The men emphasised that the care pathway should not be prolonged. This pre-trial feedback from patient and public representatives helped develop the initial trial design.

Members of the same panel also contributed to the design of the letter of invitation to participate, and the patient information sheet, as well as contributing during the recruitment period of the trial. We sought insight about ways to improve the trial experience for men, resulting in changes to documentation and processes (e.g. consent and questionnaire completion methods; see *Amendments*). There was also direct involvement in an early press release, which aimed to increase the trial profile.

Throughout the trial, there was a continued presence of two patient representatives at TMG and TSC meetings. They advised on trial progress, participant retention, newsletters (i.e. content, design and frequency) and continued development, as well as assisting with ideas surrounding the reporting and dissemination of findings.

Participants

Urological departments from 26 NHS hospitals across England recruited men from October 2014 to December 2016.

Inclusion criteria
Given the pragmatic nature of this trial, eligibility criteria were minimal and consisted of recruiting men (aged $\geq$ 18 years) who were seeking further treatment, which may include surgery, for their bothersome LUTS. Men who were already on surgery waiting lists were also eligible, providing they were willing to have UDS first, if assigned to that arm, and to reconsider treatment options if deemed appropriate following the UDS assessment.

Exclusion criteria
Men were not eligible if they met any of the following exclusion criteria:

- inability to pass urine without a catheter [excluding clean intermittent self-catheterisation (ISC) after passing urine, in order to complete emptying of their bladder]
- have a relevant neurological disease (such as stroke, multiple sclerosis, Parkinson's disease or spina bifida)
- currently undergoing treatment for prostate or bladder cancer
- previously had prostate surgery
- not medically fit for surgery
- do not consent to being randomised or comply with essential trial procedures.

Interventions

Men were randomised to a diagnostic pathway based on routine care (i.e. assessment as set out in the NICE clinical guidance on male LUTS:[5] routine care control arm) or a pathway that included UDS (routine care plus UDS: intervention arm). UDS is a diagnostic test to evaluate bladder function, assessing how well a person's bladder stores urine and how well they pass urine (voiding). The test used catheters to measure bladder and abdominal pressures during bladder filling and passing urine. Any change in abdominal pressure was also detected in the bladder and a computer calculated the difference between bladder and abdominal pressure throughout the test. Detrusor pressure was then derived and used to assess voiding function and urine storage. This knowledge was then used to detect whether or not BOO is present. Once voiding function was tested, an opinion on surgery effectiveness in treating the man could be made.

Outcome measures

Primary outcome

The IPSS, a well-established and validated patient-reported outcome, was collected at baseline and at 6, 12 and 18 months after randomisation. Men filled in a questionnaire concerning their LUTS, which produced a score from 0 to 35 (with higher scores indicating more severe symptoms).

Secondary outcomes

Key secondary (surgery)

The number of men having surgery within 18 months was recorded in each arm and the proportion of men having surgery was calculated using the number of men with a completed 18-month case report form (CRF) as the denominator. Men who withdrew consent for a notes review to be carried out at 18 months, or who died, were removed for this outcome. As a pragmatic trial, standard practice for the centres were followed, relating to the type of surgery, LUTS medications or other treatments.

Adverse events

The number of AEs were recorded in each arm, as well as the severity, expectedness and relationship to testing and treatment. *Box 2* provides the definition of a serious adverse event (SAE) and lists expected and related AEs. When events were related to surgery, they were given a Clavien–Dindo classification,[23] of which there are five grades (plus two subgroups for grades 3 and 4); see Dindo *et al.*[23] for further details.

BOX 2 Definition of a SAE and lists of expected and related events for diagnostic (testing) and prostate surgery procedures

---

Serious adverse event

An AE was defined as serious if one of more of the following applied to the AE:

- resulted in the death of the participant
- was life-threatening (i.e. an event whereby the participant was at risk of death at the time of the event; it does not refer to an event that, hypothetically, may have caused death if it were more severe)
- required hospitalisation or prolongation of existing inpatient hospitalisation
- resulted in persistent/significant disability/incapacity
- was considered medically significant by the investigator (i.e. important AEs that were not immediately life-threatening or did not result in death or hospitalisation, but may have jeopardised the subject or required intervention to prevent one of the other outcomes listed above, may also be considered serious; medical judgement was exercised in deciding whether or not an AE is serious in other situations).

Expected, related adverse events during/after any diagnostic procedures

- UTI.
- Bacteriuria.
- Haematuria.
- Urinary retention.
- Discomfort.
- Dysuria.
- Urethral trauma.

BOX 2 Definition of a SAE and lists of expected and related events for diagnostic (testing) and prostate surgery procedures (adapted from Bailey *et al.*[19]) (*continued*)

Expected, related adverse events for prostate surgery

- Excess blood loss (of > 500 ml).
- Blood transfusion.
- Urethral injury.
- Bladder injury.
- Bowel injury.
- Injury to blood vessels or nerves.
- Anaesthetic complications.
- Thrombosis/deep-vein thrombosis/pulmonary embolism.
- Prolongation of postoperative catheterisation.
- Recatheterisation.
- UTI.
- Other infection (sepsis, septicaemia, abscess).
- New urinary tract symptoms.
- Constipation.
- Discomfort/pain.
- New sexual problems.
- Death.

Reproduced from Bailey *et al.*[19] This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

All AEs (serious and non-serious) were reviewed by an independent clinician, including the relatedness to testing and treatment and Clavien–Dindo scoring for surgery-related events, to ensure uniformity in the classifications across the centres and check for bias in the classifications. Any disagreements with the assignment or classification were revised when appropriate.

Patient-reported outcome measures: ICIQ-MLUTS

In the 18-month questionnaire, men were asked to fill out the International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms (ICIQ-MLUTS). [Note that copies of the International Consultation on Incontinence Questionnaires can be requested via the website, URL: www.iciq.net (accessed 29 May 2019).] Both a voiding and incontinence score were generated from this questionnaire as well daytime and night-time voiding frequency data. This information was also collected at baseline and at 6 and 12 months.

Patient-reported outcome measures: ICIQ-MLUTSsex

In the 18-month questionnaire, men were asked to fill out the International Consultation on Incontinence Questionnaire – Sexual Matters associated with Male Lower Urinary Tract Symptoms (ICIQ-MLUTSsex). The ICIQ-MLUTSsex contained data on erection and ejaculation quality and how bothersome this was for them. This information was also collected at baseline and at 6 and 12 months.

International Prostate Symptom Score quality of life

As part of the IPSS questionnaire, men were asked how they would feel if they were to spend the rest of their life with their urinary condition, rated on a scale from 0, 'delighted', to 6, 'terrible'.

Maximum urinary flow rate

Uroflowmetry, evaluating the Q_{max}, VV and PVR, was measured as part of the 18-month clinic. When men did not turn up to the clinic, this measure was collected using case note reviews and the measurements closest to 18 months after randomisation were utilised. If two measures were taken on the same day, the higher flow rate was used. Uroflowmetry data (Q_{max}, VV and PVR) were also collected at baseline; if men had recently had a uroflowmetry test prior to joining the trial, these data were used to avoid unnecessary repetition for the patient. Additional uroflowmetry data were collected for men undergoing surgery in both arms, approximately 4 months after surgery ($\pm$ 1 month).

Patient-reported outcome measures: ICIQ-UDS-S

Overall satisfaction with the UDS assessment was captured using the International Consultation on Incontinence Questionnaire – urodynamics satisfaction (ICIQ-UDS-S), which asked about the details of the procedure, patient satisfaction and whether or not they would recommend the test. Although not a testable outcome, the satisfaction with the UDS procedure was explored descriptively.

Non-inferiority margin

A non-inferiority design was chosen to assess whether or not men randomised to receive UDS would have patient-reported urinary symptoms that were better than routine care or no worse than an acceptable number of symptoms. The sample size calculation was based on the consideration that men randomised to UDS should have symptoms that are non-inferior to those who are randomised to routine care. For the primary outcome, a lower IPSS in the intervention arm, or of no more than 1 point higher, was hypothesised as suggesting non-inferiority. The team felt that these were appropriate for the following reasons:

- The minimally clinically important difference for the IPSS is generally accepted to be a 3-point difference;[24] however, given that a difference smaller than this may involve a substantial difference on an individual subscale, it was felt that the non-inferiority margin should reflect a more sensitive change.
- A single void per night has shown to not be too much of a problem for men suffering with nocturia.[25] However, at least two voids have been shown to be substantially bothersome; therefore, a 1-point difference would be able to detect this subtle but pivotal change.
- A difference of 1 point is much more conservative than a difference of 2 or 3 points, requiring a much larger sample size, and would therefore reduce the risk of falsely claiming non-inferiority.

Sample size

The sample size calculation was based on the non-inferiority primary outcome of IPSS at 18 months after randomisation, as well as the key secondary outcome of surgery rate. The non-inferiority margin was set at 1 point and a one-sided *t*-test with common standard deviation (SD) of 5, 80% power and 5% alpha gave a required sample size of 310 participants per arm. It was estimated that 20% of participants would be lost to follow-up or withdraw (attrition); therefore, the sample size was inflated to 388 men per arm (776 men in total).

The key secondary outcome was based on a superiority outcome of surgery, for which it was anticipated that the UDS procedure may reduce the proportion of men having surgery. Using hospital audit data from Bristol, for 5670 men presenting with LUTS suggestive of poor or obstructive urine flow,[19] the team sought to find an absolute difference of 13% based on 73% of men in the routine care arm having surgery compared with 60% in the UDS arm. To achieve 90% power with a two-sided alpha of 5% required a sample size of 291 men per group. Inflating for anticipated attrition rates led to a sample size of 364 men per arm, smaller than for the primary outcome. Therefore, the team aimed for a recruitment total of 800 men.

Randomisation and implementation

Eligible men were randomised using 'simple randomisation', whereby trained staff (e.g. research nurse, administrators) utilised a telephone- and web-based randomisation system that randomised men to UDS or routine care, with no stratification or minimisation techniques. Randomisation was carried out in the Bristol Randomised Trials Collaboration (BRTC). Originally, the trial team planned to stratify the randomisation by centre, but, given the use of simple randomisation, this was not the case. It was, however, adjusted for in all analyses.

Blinding

Given the nature of UDS, men were not blinded to their trial arm. The trial manager and administrative staff, although unblinded to enable individual data collection and AE reporting, were blinded to aggregate data. Centre (hospital) staff had access to unblinded data to enable them to make clinical decisions about the management of a man's LUTS, including the appropriateness of surgery.

All investigators remained blinded to aggregate data throughout recruitment and analysis. The senior statistician (PSB) had not seen any data when writing the SAP and remained blinded until the analysis had been finalised. The junior statistician (GJY) had unblinded access to the data to report safety and outcome data to the DMC. The protocol was written before recruitment ended and published in a peer-reviewed journal.[19] The SAP was written and agreed by the trial team, DMC and TSC in September 2016 prior to recruitment end. It was submitted for publication on the 14 December 2016 and underwent minor revisions before being published on 3 October 2017.[20]

Data collection

Data were collected at certain points from various data collection forms (*Table 2*). The primary outcome (IPSS) was collected at baseline and at 6, 12 and 18 months after randomisation. The ICIQ-MLUTS and ICIQ-MLUTSsex questions were also asked at these time points. A 3-day bladder diary was distributed at baseline and 18 months; and will be analysed separately at a later date as an exploratory analysis.

TABLE 2 Data collection in UPSTREAM (adapted from Bailey *et al.*[19])

Data collection tool	Baseline	UDS only After	Surgery only Perioperative	4 months[a] after	Follow-up 6 months	12 months	18 months
CRF	✓	✓	✓	✓			✓
IPSS	✓				✓	✓	✓
ICIQ-MLUTS	✓				✓	✓	✓
ICIQ-MLUTSsex	✓				✓	✓	✓
ICIQ-UDS-S		✓					
Flow rate/PVR	✓			✓			✓
Bladder diary	✓						✓
Case note review							✓

a ± 1 month.

As some CRFs and questionnaires were completed earlier/later than their scheduled time in the follow-up, a post hoc sensitivity analysis was added that included questionnaires only from within a specific time frame (prior to analysis, the team established a time frame for which questionnaires would be included in this sensitivity analysis). An 18-month questionnaire would be accepted only if it was within the window of 16–20 months from the date of randomisation. The baseline questionnaire, which was adjusted for, was accepted only if it was within 6 months of the randomisation date.

Statistical methods

The main statistical analyses were prespecified using a SAP, which was completed prior to recruitment end.[20] The database was closed on the 20 August; final analysis started soon after and finished in October 2018. Stata® 15.1 (StataCorp LP, College Station, TX, USA) was used for all statistical analyses in this trial. Binary outcomes were presented as n (%), whereas continuous outcomes were presented as mean (SD) or median (interquartile range), as appropriate. Any deviations from these methods were considered 'post hoc' analyses and clearly labelled as such.

Primary analysis

The IPSS was collected from the 18-month questionnaire. The primary analyses were conducted under the intention-to-treat (ITT) principle using multivariable linear regression. Both centre and the baseline IPSS were adjusted for in the primary analysis. Results were based on the non-inferiority margin prespecified in the trial design process. As the primary analysis tested non-inferiority, interpretation of the primary analysis results focused on observed differences, and 95% confidence intervals (CIs) for the between-group comparisons. When CIs lay below the non-equivalence margin, the two arms were deemed to be equivalent.

H_0: routine care leads to IPSSs that are at least 1 point lower, on average, than in the UDS arm (higher IPSSs signify a poorer outcome).

H_1: UDS is non-inferior to routine care, with IPSSs that are lower than in the routine care arm or less than 1 point higher.

Secondary analyses

In a similar way to the primary analysis, all secondary analyses were conducted using ITT and adjusting for centre and baseline measures (when appropriate). However, all secondary analyses tested for superiority as opposed to non-inferiority.

The key secondary outcome was surgery rates at 18 months. Using logistic regression, the proportion of men receiving surgery for BOO was compared between the two arms. The denominator for this outcome was the number of men who were followed up at 18 months, either by clinic appointment or by case notes review. The proportion having surgery by their 18-month follow-up was recorded via perioperative CRFs and 18-month CRFs. If men withdrew before their 18-month time point, or did not attend their clinic appointment, details were obtained via notes review, when permissible. We also explored whether or not men agreed with their doctor's recommendation on listing for surgery.

Adverse events were collected throughout each man's 18-month period of follow-up. The relationship to diagnostic testing and/or treatment was considered, as well as the severity and expectedness. When events were deemed related to surgery, a Clavien–Dindo score[23] was assigned. The number of deaths and the number of acute urinary retention cases were also recorded separately in each arm. Events were compared using logistic and ordinal logistic regression.

Each of the following patient-reported outcome measures (PROMs) were compared between the arms at 18 months, using linear and logistic regression:

- IPSS QoL
- ICIQ-MLUTS (voiding scale, incontinence scale, daytime frequency and nocturia)
- ICIQ-MLUTSsex (erection and ejaculation quality, painful ejaculation and overall effect on sex life).

The following dichotomous variables were created from the ICIQ-MLUTS for ease of reporting and interpretation:

- Daytime frequency (eight or more times): question 13a, coded as 1 if the man ticked '9 or 10 times', '11 or 12 times' or '13 or more times'.
- Nocturia (one or more times per night): question 14a, coded as 1 if the man ticked 'two', 'three' or 'four or more'.

Dichotomous variables were also created from the ICIQ-MLUTSsex:

- Erections (reduced or none): question 2a, coded as 1 if the man ticked 'yes, with reduced rigidity', 'yes, with severely reduced rigidity' or 'no, erection not possible'.
- Ejaculation (reduced or none): question 3a, coded as 1 if the man ticked 'yes, reduced quantity', 'yes, significantly reduced quantity' or 'no ejaculation'.
- Painful ejaculation (yes): question 4a, coded as 1 if the man ticked 'yes, slight pain/discomfort', 'yes, moderate pain/discomfort' or 'yes, severe pain/discomfort'.
- Urinary symptoms affected sex life: question 5a, coded as 1 if man ticked 'a little', 'somewhat' or 'a lot'.

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

Maximum urinary flow rate was collected at 18 months, either at the 18-month clinic visit or via case note review. When multiple measures were recorded, the measure closest to the 18-month time point (548 days after randomisation) was used. Linear regression was used to compare flow rates between the arms, adjusting for baseline urinary flow rate and centre.

Satisfaction with UDS was explored descriptively as it was collected only for those men who had UDS.

Additional cost-effectiveness and qualitative analysis outcomes are described and presented in *Chapters 4* and *5*.

Sensitivity analyses
Several prespecified sensitivity analyses were conducted to test the robustness of the results from the statistical analyses to increase understanding of the relationship between the dependent and independent variables for the primary analysis and, in some circumstances, the key secondary analysis:

Per-protocol analysis
The primary and key secondary analyses were repeated, analysing men who received UDS, having been allocated to it, against those who did not receive it, having been allocated to routine care. Therefore, any men who had not complied with their randomised pathway were excluded from the analyses.

Complier-average causal effect analysis
The complier-average causal effect (CACE) analysis used the data from those men who did receive their randomised treatment but, unlike the primary analysis, incorporated the treatment-received variable as the independent variable and randomisation as an instrumental variable, using the 'iv regress sls' command in Stata.

Mixed-effects model

This analysis allowed us to look at the total 18-month impact of the intervention and whether or not there was a difference between the two arms. The mixed-effects repeated measures approach was adopted to assess the total 18-month impact, while accounting for missing data at each 6-month time point.

Imputation using 6- and 12-month data

When baseline IPSSs were missing, the 6-month scores were utilised if no intervention procedures had taken place (e.g. surgery). When 18-month scores were missing, the 12-months scores were utilised if all intervention procedures had taken place by this time point, if they were scheduled to take place.

Imputation for missing data

Men included in the analysis were compared with those men who were not followed up at 18 months. The missingness mechanism was assessed and multiple imputation methods were adopted. The randomisation seed was prespecified (648) to allow reproducible results and reduce any tampering.

Adjustment for clinically important confounders

When the trial was designed, the clinicians produced a list of clinically important confounders that should be considered in the analysis of the primary outcome. These were centre, age, comorbidities and symptom severity. Therefore, these were adjusted for in a sensitivity analysis.

Adjustment for imbalance at baseline

A prespecified measure of imbalance was used to compare the arms and any continuous measures that were at least 0.5 SDs apart, or had an absolute difference of 10% (for binary and categorical outcomes), were adjusted for.

Adjustment for time from surgery

When UPSTREAM was designed, the team envisaged that the 18-month follow-up would incorporate a 6-month post-surgery gap that would allow any side effects of treatment to subside. It became apparent that assessment and treatment pathways across the 26 hospitals varied, and waiting lists for surgery were longer than anticipated, potentially influencing the 18-month symptom scores. Therefore, the time between surgery and the 18-month questionnaire was calculated, imputing a time of 1000 days when no surgery occurred.

When the 18-month time points were scrutinised during follow-up, the team chose to conduct a post hoc sensitivity analysis that excluded follow-up outside a certain trial window (16–20 months). This was done to ensure that the results we achieved were truly reflective of an 18-month follow-up.

Subgroup analyses

Prespecified subgroups were used to test whether or not the differences between the two arms were more pronounced in certain subgroups of men. Although underpowered, tests of interaction between the dichotomised/categorical variables and trial arm were carried out to test whether or not the treatment effect differed between subgroups. These interaction terms were added to the primary analysis model.

Subgroup analyses included:

- age (split by median age)
- flow rate (> 12 ml/second vs. ≤ 12 ml/second)
- maximum VV (< 200 ml vs. ≥ 200 ml)
- storage dysfunction/nocturia (yes vs. no)
- severity of storage LUTS, questions 2, 4 and 7 in the IPSS questionnaire (split by median).

Assessing non-inferiority

Although a *p*-value can be very helpful in determining superiority, it cannot determine non-inferiority or equivalence. Therefore, for the primary outcome of this trial, emphasis was placed on the 95% CIs and their positioning around the non-inferiority margin. The variety of potential conclusions are demonstrated in *Figure 2* (adapted from Schumi and Wittes[26]).

Future analysis plans

A bladder diary analysis plan has been prespecified, which will look at clinical characteristics of the men, such as average sensation score, total urgency and frequency score, and VV. It will assess the quality of the bladder diary by comparing it with the measures from the trial, for example VV, as well as measures from the validated questionnaires (e.g. IDUF and nocturia). By looking at the daytime frequency and nocturia, we can also assess the International Consultation on Incontinence Questionnaire and IPSS to see whether or not one is more closely related to the results from the bladder diary.

The treatment pathways are of interest to the trial team, so it is probable that they will go through each of the men's treatments, both surgical and conservative, to gain an understanding of a typical pathway for the treatment of LUTS.

UPSTREAM has been awarded further follow-up funding to assess the longer-term impact of LUTS and its treatment.
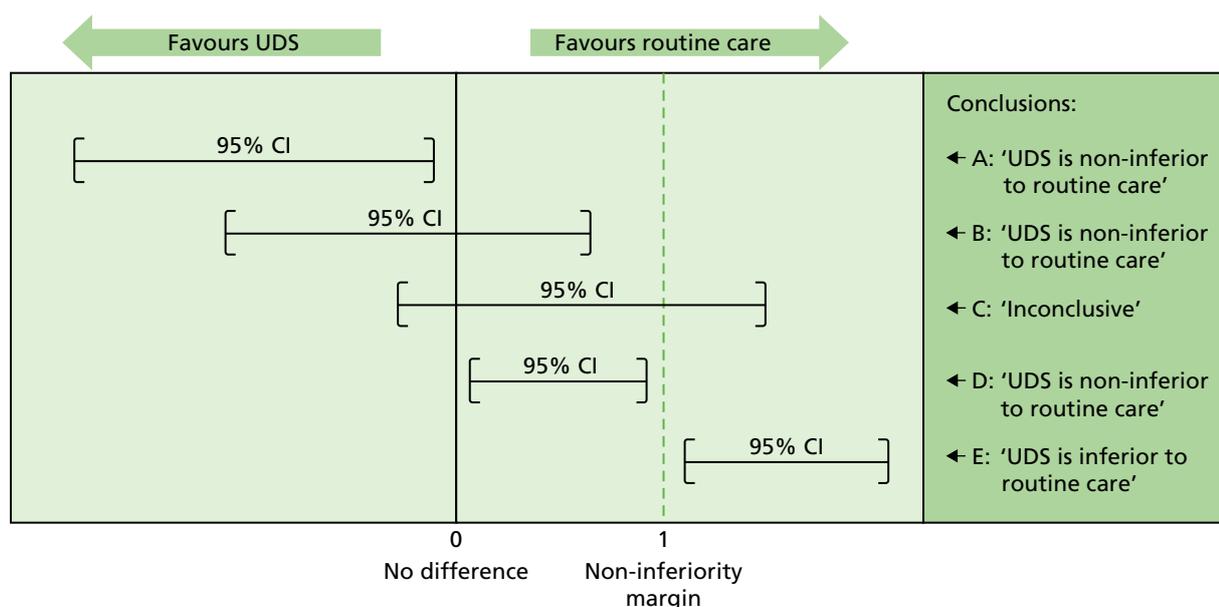
FIGURE 2 Assessing the potential non-inferiority conclusions for UPSTREAM. Treatment difference (UDS – routine care). Adapted from Schumi and Wittes[26] © 2011 Schumi and Wittes; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chapter 3 Results

Some of the work reported in this chapter may overlap with content published in dedicated journal papers, which we wish to acknowledge.[27–29]

Some text has been based on Drake et al.[28] Reprinted from *Eur Urol*, 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 [published online ahead of print June 30 2020], Copyright 2020, with permission from Elsevier. Some text has also been based on Aiello et al.[29] This is an open access article under the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons. org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Participant flow

Figure 3 shows the layout of the trial and the different levels of dropout and analysis. Overall, 67 men withdrew from the trial within 548 days (18 months) of randomisation (seven of whom requested complete data withdrawal), with a median withdrawal time of 8 months (of 60 men with an available date). A further nine men from each arm withdrew after 548 days; however, these men were not included in the Consolidated Standards of Reporting Trials (CONSORT) flow diagram. There were 18 withdrawals in the UDS arm between randomisation and 6 months, but only six withdrawals in the routine care arm. When looking at the reasons for withdrawal, two did not like their randomised allocation and six did not give an explanation; neither of these reasons was given in the routine care arm. There were nine deaths in the UDS arm, compared with only two in the routine care arm; all were unrelated to trial procedures (see *Table 8*). We were also made aware of four unrelated deaths that occurred after men withdrew (two in the UDS arm and two in the routine care arm), but no further details were given. For the IPSS, 80% and 84% of men had available data at 18 months, in the UDS and routine care arms, respectively. After adjustment for centre and baseline IPSS, 328 and 313 men were analysable at 18 months for the UDS and routine care arms, respectively. The sample size calculation had estimated that 310 men per group would provide 80% power to answer the primary research question.

Recruitment

Of the 8671 men screened for eligibility, 1482 (17%) were considered eligible, of whom 820 (55%) were randomised to UPSTREAM (427 men in the UDS arm and 393 men in the routine care arm). Further details of patient screening are provided in *Appendix 1* (see *Tables 34–36*). The first man was randomised in October 2014 and the last man in December 2016. There was no obvious seasonal variation; however, recruitment was slower during school holidays, when fewer recruiting staff were available (e.g. August, September, December) (*Figure 4*).

Baseline data

Baseline comparisons for UPSTREAM are given in *Tables 3* and *4*. The team prespecified in the analysis plan[20] that any baseline characteristics that differed by > 10% (categorical variables) or by more than 0.5 SDs (continuous variables) would be adjusted for in a sensitivity analysis. None of the baseline characteristics met these criteria, so the sensitivity analysis was not carried out. The two arms were well balanced with respect to baseline characteristics, suggesting that allocation concealment in randomisation was successful.
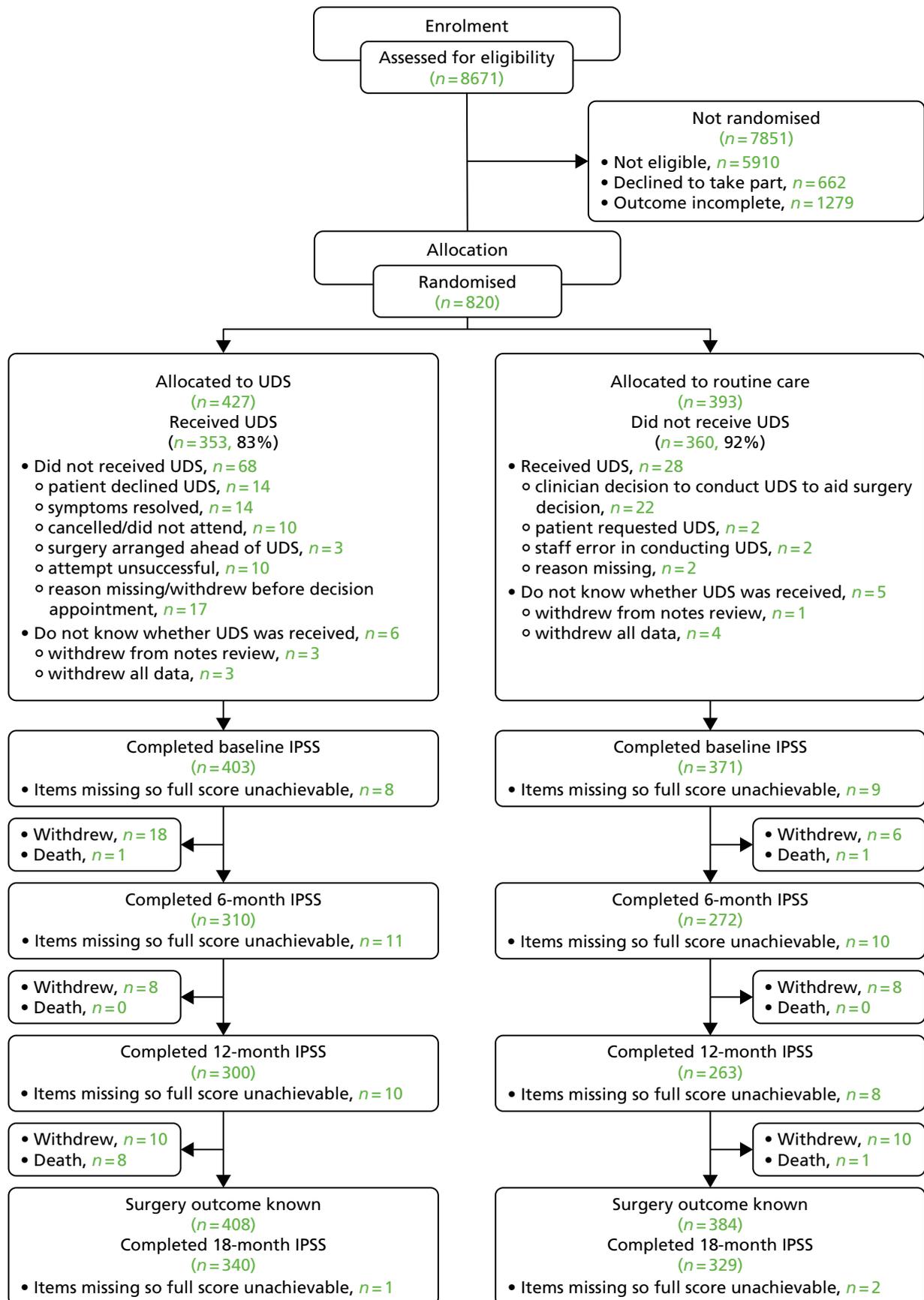
FIGURE 3 The UPSTREAM trial Consolidated Standards of Reporting Trials (CONSORT) flow diagram.
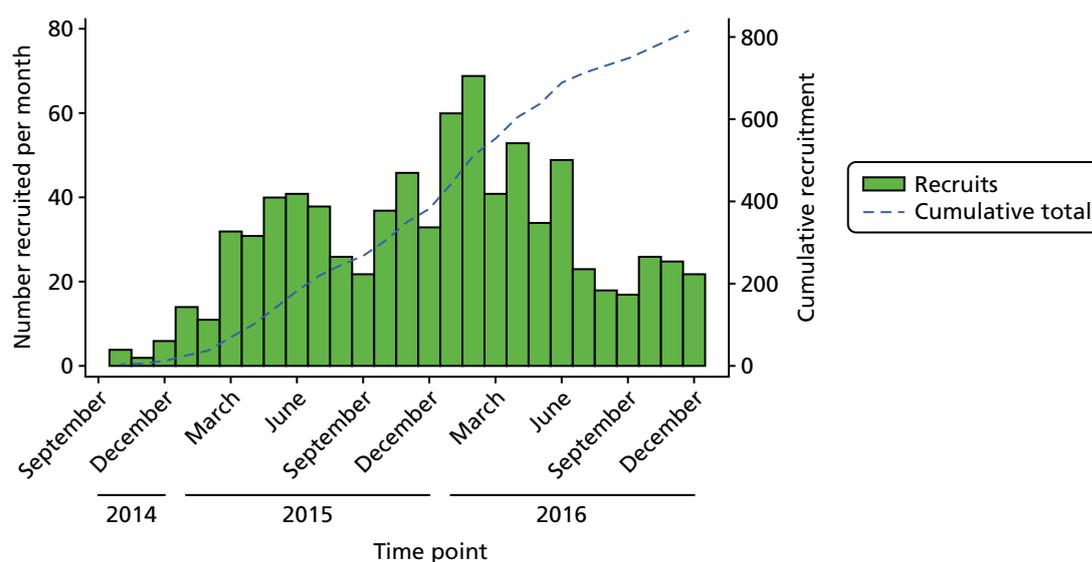
FIGURE 4 The UPSTREAM recruitment chart.

TABLE 3 Baseline sociodemographics of eligible men who underwent randomisation

Baseline sociodemographics	N^a	UDS	N^a	Routine care
Total number of participants	427		393	
Age (years), mean (SD)	424	67.51 (9.59)	389	67.81 (8.79)
Centre,[b] n (%)				
1		56 (13)		58 (15)
2		12 (3)		20 (5)
3		35 (8)		27 (7)
4		29 (7)		26 (7)
5		29 (7)		18 (5)
6		5 (1)		4 (1)
7		17 (4)		9 (2)
8		9 (2)		12 (3)
9		17 (4)		14 (4)
10		11 (3)		6 (2)
11		8 (2)		7 (2)
12		16 (4)		17 (4)
13	424	15 (4)	393	11 (3)
14		17 (4)		19 (5)
15		14 (3)		13 (3)
16		21 (5)		23 (6)
17		24 (6)		13 (3)
18		14 (3)		12 (3)
19		3 (1)		6 (2)
20		15 (4)		16 (4)
21		13 (3)		13 (3)
22		9 (2)		7 (2)

continued

TABLE 3 Baseline sociodemographics of eligible men who underwent randomisation (*continued*)

Baseline sociodemographics	*N*[a]	UDS	*N*[a]	Routine care
23		13 (3)		21 (5)
24		10 (2)		9 (2)
25		11 (3)		9 (2)
26		4 (1)		3 (1)
Ethnicity, *n* (%)				
White		377 (91)		356 (93)
Black/African/Caribbean/black British		8 (2)		6 (2)
Mixed/multiple ethnic groups	415	17 (4)	383	11 (3)
Asian/Asian British		2 (< 1)		1 (< 1)
Other ethnic group		3 (1)		2 (1)
Disclosure declined		8 (2)		7 (2)
IMD scores 2015 (based on postcodes)				
IMD score 2015,[c] median (IQR)	411	14 (8–22)	383	14 (8–24)
Quintile 1 (most deprived), *n* (%)		43 (10)		61 (16)
Quintile 2, *n* (%)		75 (18)		49 (13)
Quintile 3, *n* (%)		92 (22)		91 (24)
Quintile 4, *n* (%)		106 (26)		86 (22)
Quintile 5 (least deprived), *n* (%)		95 (23)		96 (25)

IMD, Index of Multiple Deprivation.
a   The number of men who we have data for (denominator): three men in the UDS arm and four men in the routine care arm requested that all of their data be withdrawn; therefore, the maximum values are 424 and 389, respectively (apart from centre).
b   Centre names have been replaced with numeric identifiers for the purpose of reporting.
c   Higher scores mean higher levels of deprivation. As data were skewed for these outcomes, medians and IQRs are presented.[30]
Reprinted from *European Urology Focus*, vol. 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). pp. 340–50, copyright 2019,[27] with permission from Elsevier.

TABLE 4 Baseline clinical characteristics of eligible men who underwent randomisation

Clinical baseline characteristic	*N*[a]	UDS	*N*[a]	Routine care
Comorbidities at baseline, *n* (%)	420	281 (67)	383	260 (68)
DRE findings,[b] *n* (%)				
No abnormality	395	108 (27)	375	120 (32)
Benign enlargement	395	312 (79)	375	287 (77)
Suspected prostate cancer	395	16 (4)	375	8 (2)
Other	395	22 (6)	375	20 (5)
Uroflowmetry,[c] median (IQR)				
Q_{max} (ml/second)	402	10.20 (7.40–15.00)	371	11.00 (7.90–58.30)
PVR (ml)	401	100.00 (40.00–180.00)	373	100.00 (45.00–189.00)
VV (ml)	405	215.00 (133.00–318.00)	376	214.00 (149.50–316.00)

TABLE 4 Baseline clinical characteristics of eligible men who underwent randomisation (*continued*)

Clinical baseline characteristic	N[a]	UDS	N[a]	Routine care
Additional (discretionary) tests, *n* (%)				
PSA test		57 (14)		57 (15)
Cystoscopy		44 (11)		25 (7)
Urinalysis		59 (14)		59 (15)
Urea and electrolytes	413	18 (4)	383	17 (4)
Kidney ultrasound		14 (3)		11 (3)
Voiding urology cytology		2 (< 1)		2 (1)
Prostate volume measurement		15 (4)		7 (2)
IPSS: symptom severity at baseline, mean (SD)				
Total IPSS	403	18.52 (6.90)	371	19.39 (7.14)
Incomplete emptying	411	2.64 (1.71)	379	2.88 (1.72)
Frequency	411	3.36 (1.35)	379	3.56 (1.30)
Intermittency	411	2.58 (1.69)	379	2.65 (1.62)
Urgency	409	2.60 (1.68)	379	2.80 (1.66)
Weak stream	409	3.17 (1.57)	379	3.16 (1.61)
Straining	408	1.56 (1.56)	377	1.67 (1.66)
Nocturia	410	2.60 (1.32)	379	2.72 (1.28)
IPSS QoL	411	4.07 (1.36)	379	4.20 (1.25)
ICIQ-MLUTS				
Voiding score,[d] mean (SD)	394	8.88 (4.04)	370	9.30 (4.38)
Incontinence score,[e] mean (SD)	395	5.01 (3.37)	369	5.19 (3.27)
Daytime frequency (eight or more times), *n* (%)	398	160 (40)	374	169 (45)
Nocturia (one or more times per night), *n* (%)	398	300 (75)	374	301 (80)
ICIQ-MLUTSsex, *n* (%)				
Erections (reduced or none)	389	277 (71)	362	275 (76)
Ejaculation (reduced or none)	383	300 (78)	359	295 (82)
Painful ejaculation (yes)	359	56 (16)	343	71 (21)
Urinary symptoms affected sex life?	378	259 (69)	358	233 (65)

IQR, interquartile range.
a The number of men who we have data for (denominator): three men in the UDS arm and four men in the routine care arm requested that all of their data be withdrawn; therefore, the maximum values are 424 and 389, respectively (apart from centre).
b 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 answered yes/no to at least one finding.
c As data were skewed for these outcomes, medians and IQRs are presented.
d On a scale of 0–20, with larger scores indicating more severe symptoms.
e On a scale of 0–24, with larger scores indicating more severe symptoms.
Reprinted from *European Urology Focus*, vol. 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). pp. 340–50, copyright 2019,[27] with permission from Elsevier.

Although stratified randomisation (by centre) was planned, simple randomisation was employed. It was observed mid-trial that more men were being recruited to the UDS arm, and this was not equally distributed across centres. Although this could not be adjusted, examination of the randomisation system confirmed that it was functioning correctly and that any imbalance was due to chance and not significant. This was also reviewed by the TSC, which agreed with this conclusion.

When comparing the characteristics of those men analysed at 18 months with those who withdrew or were lost to follow-up (not analysed), we see very similar characteristics (*Table 5*). There were more men with comorbidities among those not analysed (71%), which was not surprising given that that this was one of the main reasons for withdrawal. The only baseline measures that differed by the prespecified 10% and 0.5 SDs were the proportions of men who had a cystoscopy or PSA test at baseline. Of the 35 men who had a PSA test, 15 withdrew (seven because of poor health). Of the 26 men who had a cystoscopy, five withdrew because of poor health and two died.

TABLE 5 Baseline characteristics for those men included and those not included in the primary outcome

Baseline characteristics	N^a	Analysed at 18 months	N^a	Not analysed
Age (years), mean (SD)	641	67.79 (8.46)	172	67.14 (11.62)
Clinical baseline characteristic, *n* (%)				
Comorbidities	633	421 (67)	170	120 (71%)
Additional tests, *n* (%)				
PSA test	639	79 (12)[b]	157	35 (22)[b]
Cystoscopy	639	42 (7)[b]	157	27 (17)[b]
Urinalysis	639	85 (13)	157	33 (21)
Kidney ultrasound	639	15 (2)	157	10 (6)
Cytology	639	4 (1)	157	0 (0)
Prostate volume measurement	639	14 (2)	157	8 (5)
Urea and electrolytes	639	24 (4)	157	11 (7)
IPSS: symptom severity at baseline, mean (SD)				
Total IPSS	641	18.83 (7.05)	133	19.47 (6.89)
IPSS QoL	640	4.09 (1.31)	150	4.31 (1.30)
ICIQ-MLUTS				
Voiding score,[c] mean (SD)	621	9.10 (4.15)	143	9.03 (4.49)
Incontinence score,[d] mean (SD)	622	5.01 (3.27)	142	5.47 (3.47)
Daytime frequency (eight or more times), *n* (%)	626	262 (42)	146	67 (46)
Nocturia (one or more times per night), *n* (%)	626	484 (77)	146	117 (80)
ICIQ-MLUTSsex, *n* (%)				
Erections (reduced or none)	613	449 (73)	138	103 (75)
Ejaculation (reduced or none)	605	489 (81)	137	106 (78)
Painful ejaculation (yes)	574	106 (18)	128	21 (16)
Urinary symptoms affected sex life?	600	397 (66)	136	95 (70)

a The number of men for whom we have data (denominator): three men in the UDS arm and four men in the routine care arm requested that all of their data be withdrawn; therefore, the maximum values are 424 and 389, respectively (apart from centre).
b Differed by at least the prespecified 10%/0.5 SDs.
c On a scale of 0–20, with larger scores indicating more severe symptoms.
d On a scale of 0–24, with larger scores indicating more severe symptoms.

Numbers analysed

The numbers who had outcome data were also relatively balanced between the groups; however, the number adhering to the diagnostic pathway (i.e. trial arm) to which they were assigned differed; $p < 0.001$ (*Table 6*). A larger proportion of men in the routine care arm adhered to the care pathway (93%) than those randomised to UDS (84%), with reasons for non-adherence shown in *Figure 3*. As also seen in *Figure 3*, the number of withdrawals within 548 days was 67 and the number of deaths was 11.

For the surgery rate outcome, the following men were removed from the final analysis: 7 who completely withdrew; 11 men who died prior to their 18-month follow-up time point; and 10 men who withdrew before their 18-month follow-up time point and declined permission to obtain a notes review. Owing to the higher number of deaths in the UDS arm ($p = 0.047$), a lower number of men were analysable for surgery rate (96% vs. 98% for the UDS and routine care arms, respectively).

Reasons for withdrawal were relatively balanced between the arms. There were more deaths in the UDS arm, but these did not appear to be related to the intervention, or any UPSTREAM pathway procedure/treatment; therefore, this appears to be a chance finding (*Tables 7* and *8*). There were four men in the routine care arm who felt that there were too many questionnaires, compared with no men in the UDS arm.

Statistical outcomes and estimation

International Prostate Symptom Scores
The primary outcome of IPSS at 18 months was analysed using linear regression. The scores at baseline were normally distributed, presented in *Figure 5*. Current guidelines for the measure advise that a score of 1–7 suggests mild urinary symptoms, 8–19 suggests moderate symptoms and 20–35 suggests severe symptoms. At baseline, the largest percentage of men was in the severe symptoms category (48%), followed by the moderate category (45%), whereas only 6% were in the mild symptoms category.

When we look at the histogram of scores at 18 months, we can see a large drop in IPSSs, so the majority are now in the less severe categories. Thirty per cent of men were in the mild symptoms category, 49% in the moderate symptoms category and 21% in the severe category (*Figure 6*).

TABLE 6 Adherence to diagnostic pathway (trial arm) and missing data

	UDS, *n/N* (%)	Routine care, *n/N* (%)	*p*-value[a]
Numbers of withdrawals and losses to follow-up			
Number randomised	427	393	–
Number that adhered to their randomised allocation[b]	353/421 (84)	360/388 (93)	< 0.001
Withdrew from the trial (within 548 days)	39/427 (9)	28/393 (7)	0.294
Deaths (within 548 days)	9/427 (2)	2/393 (1)	0.047
Analysable sample			
IPSS at baseline	403/427 (94)	371/393 (94)	0.989
IPSS at 6 months	310/427 (73)	272/393 (69)	0.286
IPSS at 12 months	300/427 (70)	263/393 (67)	0.304
IPSS at 18 months	340/427 (80)	329/393 (84)	0.131
Analysable for IPSS	328/427 (77)	313/393 (80)	0.327
Analysable for surgery rate	408/427 (96)	384/393 (98)	0.089

a Chi-squared test.
b Excludes men for whom we are unsure (e.g. complete withdrawal from the trial).

TABLE 7 Reasons for withdrawal

Reason for withdrawal	UDS, *n/N* (%)	Routine care, *n/N* (%)
Too many questionnaires	0/427 (0)	4ᵃ/393 (1)
Poor health	12ᵃ/427 (3)	9/393 (2)
Relocation	1/427 (< 1)	1/393 (< 1)
Declined treatment allocation	3/427 (1)	0/393 (0)
Personal reasons	5/427 (1)	3ᵃ/393 (1)
Symptoms improved	2ᵃ/427 (< 1)	1/393 (< 1)
Deceased within 18 monthsᵇ	9/427 (2)	2/393 (< 1)
Deceased after withdrawal (that were recorded)	2/427 (< 1)	2/393 (1)
Confirmed ineligible	1/427 (< 1)	1/393 (1)
Work commitments	3/427 (1)	1/393 (< 1)
Did not want hospital involvement	2/427 (< 1)	1ᵃ/393 (0)
Centre error	1/427 (< 1)	0/393 (0)
No reason given	9ᵃ/427 (2)	7/393 (2)

a Indicates that at least one of these men withdrew all data.
b Causes of death broken down in *Table 8*.

TABLE 8 Cause of death for those men within 18 months of randomisation

Trial arm	Time to death from randomisation (months)	Cause of death
Routine care	0.62	Cardiac arrest before any trial interventions organised or carried out
Routine care	13.44	Ischaemic heart disease, peripheral vascular disease, diabetes and chronic kidney disease
UDS	12.89	Multiple organ failure
UDS	17.87	Metastatic synovial sarcoma
UDS	14.82	Metastatic oesophageal cancer
UDS	15.61	Unverified, but no trial treatment/procedures within 30 days prior to the event. Patient previously removed from surgery list because of breathlessness/bronchiectasis
UDS	12.13	Pneumonia and metastatic renal cell carcinoma
UDS	13.18	Lung cancer with lung and liver metastases
UDS	14.16	Metastatic pancreatic cancer
UDS	5.41	Subarachnoid haemorrhage
UDS	13.11	Pneumonia, with no trial-related procedures within 30 days prior to the event

The median change in IPSS was –5, with a 25th and 75th percentile of –12 and 0, respectively. There were 147 men who had an IPSS that increased over the 18 months, a negative outcome. Of these, 53% underwent UDS and only 14% underwent surgery (*Figure 7*).
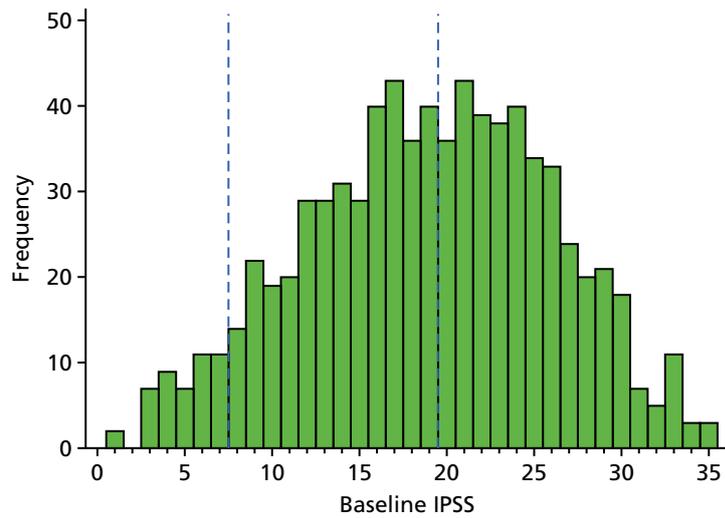
FIGURE 5 Histogram of IPSSs at baseline.
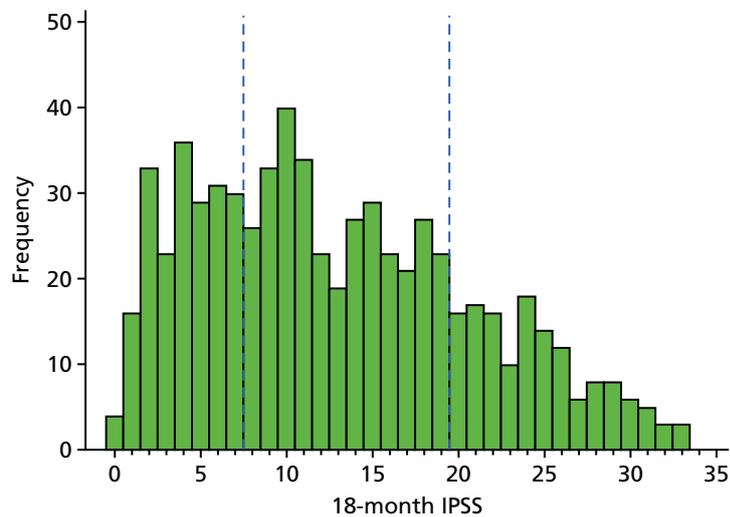
FIGURE 6 Histogram of IPSSs at 18 months.
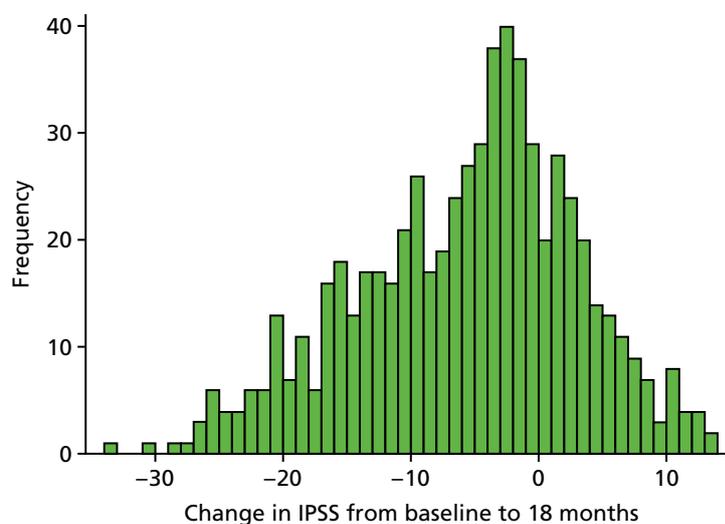
FIGURE 7 Histogram of change in IPSS.

The change in QoL score had a similar distribution to the change in overall IPSS. The median change in QoL was –1, with 25th and 75th percentiles of –3 and 0, respectively. There were 91 men who had a QoL score that increased (worsened) over the 18 months. Of these, 58% received UDS and 21% received surgery. This was explored in more detail in *Appendix 2*.

When looking at the 18-month scores across the trial arms, we see that they are very similar with respect to their distributions (*Figure 8*).

There was a similar decrease in the IPSS QoL question (on a scale of 0–6 where 0 is 'delighted' and 6 is 'terrible').

Primary analysis result

The primary analysis model included all men who answered all IPSS questions in the 18-month questionnaire and baseline questionnaire. Results are presented on an ITT basis and using complete cases. Adjusted analyses are based on a smaller sample size, as men needed to have completed both the baseline and 18-month questionnaire to be included. There were 12 and 16 men excluded from the primary analysis, from the UDS and routine care arms respectively, due to missing baseline IPSS data.

The overall SD for the 18-month IPSSs was 7.89, which was higher than the anticipated SD of 5 used in the sample size calculation. Despite this, the CI for the difference in means was still small. The non-inferiority margin was prespecified as 1 point on the IPSS scale, to demonstrate non-inferiority in the UDS arm. The lower CI level shows that, in the population, those receiving UDS have scores that are no more than 1.47 points lower than the scores of those men in the routine care arm (*Table 9*). Conversely, the upper CI level shows that those men who underwent UDS have scores that are no more than 0.80 points above the scores of those in routine care arm (a worse outcome for the UDS arm). Therefore, UPSTREAM has demonstrated that PROMs in men randomised to UDS are non-inferior to those randomised to routine care, with respect to IPSSs.

Comparing this with the scenarios explained in the methods section (see *Assessing non-inferiority*), we can see that our results conform to 'scenario B', concluding that UDS is non-inferior to routine care with regards to the IPSS (*Figure 9*).
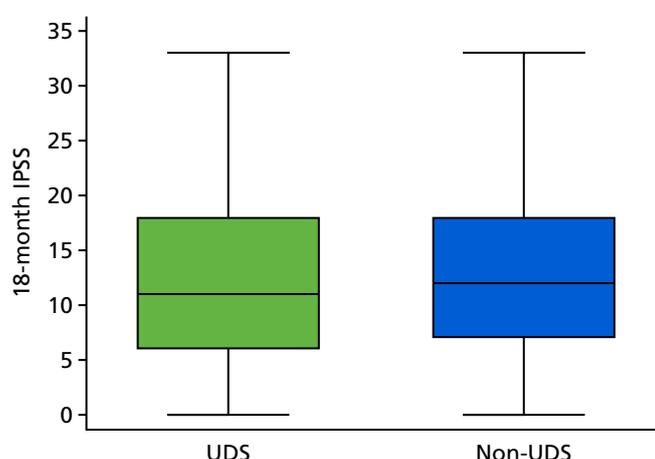
FIGURE 8 Box plot of IPSSs at 18 months, by arm.

TABLE 9 Primary analysis results

Variable	*n* (UDS : routine care)	UDS, mean (SD)	Routine care, mean (SD)	Crude difference in means (95% CI)	Adjusted difference in means[a] (95% CI)
IPSS symptom questionnaire					
Total IPSS	340 : 329	12.61 (7.92)	13.11 (7.86)	–0.49 (–1.69 to 0.70)	–0.33 (–1.47 to 0.80)
QoL score	343 : 332	2.72 (1.69)	2.74 (1.64)	–0.02 (–0.28 to 0.23)	–0.07 (–0.32 to 0.18)

a Adjusted for centre and baseline IPSS.
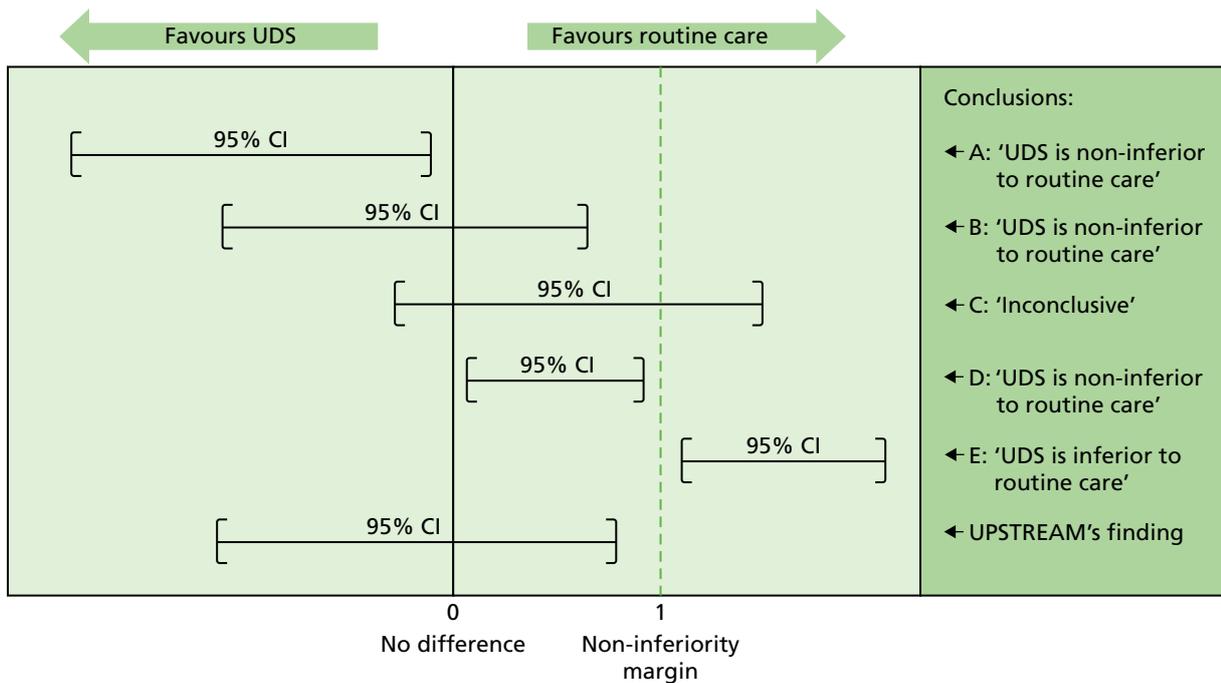
FIGURE 9 The non-inferiority conclusion for UPSTREAM. Treatment difference (UDS – routine care). Adapted from Schumi and Wittes[26] © 2011 Schumi and Wittes; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Key secondary result: surgery recommendation

In the care pathways, whether or not a man actually received surgery was as a result of several steps:

- Recommendation – whether surgery or interventional treatment was recommended by the surgeon, as opposed to behavioural therapy or medications (conservative therapy).
- Acceptance or refusal of the recommendation by the man. If the man accepted, the surgeon would then place him on the waiting list for surgery ('listing'). In the trial, the duration of wait for surgery varied between centres.
- Surgery – the 'surgery rate' key secondary outcome reflects the number of men who actually received interventional treatment, implying prior recommendation and listing.

The difference between the recommendation rate and the surgery rate reflects two main factors: (1) men who declined surgery that was recommended to them and (2) men kept on the waiting list for a long time.

The total number of men who were initially recommended surgery by their surgeon was 378 (49%) [196 men (49%) and 182 men (48%) for the UDS and routine care arms, respectively]. One centre recommended surgery to all seven of their patients (4 : 3, UDS and routine care); this was dropped from the analysis when adjusting for centre (*Table 10*).

Men did not necessarily agree with the recommendation for or against surgery and this view on agreement can also be found in *Table 10*. Recommendations were reviewed in a post hoc analysis, using criteria prespecified in consultation with urologists not taking part in UPSTREAM, before looking at the data, and agreed by the co-applicant urologists, to allow for a more uniform approach. Further details on this and the reasons for recommending surgery can be found in *Appendix 3*. The proportion of men actually listed for surgery following their decision-making appointment was 41% in both arms. Of those men listed, 264 (87%) had the surgery, 34 (11%) did not have surgery and for 5 (2%), we do not know if they received surgery (because of withdrawal/death/unable to access medical notes). Of those men initially not listed for

TABLE 10 Secondary outcome: recommendation for surgery, difference between arms

Variable	UDS, *n* (%)	Routine care, *n* (%)	OR[a] (95% CI)	*p*-value[a]
Surgeon recommendation				
Surgery	196 (49)	182 (48)	1.02 (0.76 to 1.38)	0.872
No surgery	201 (51)	196 (52)		
Surgeon recommended surgery[b]				
Patient agreement	159 (82)	153 (84)		
Patient disagreement[c]	36 (18)	29 (16)		
Surgeon did not recommend surgery[d]				
Patient agreement	200 (> 99)	190 (97)		
Patient disagreement[e]	1 (< 1)	5 (3)		
Outcome of decision-making appointment				
Listed for surgery	156 (41)	147 (41)		
Not listed for surgery	220 (59)	211 (59)		

OR, odds ratio.
a Adjusted for centre (one centre dropped because of perfect prediction).
b Patient agreement missing for one man in the UDS arm.
c 12/62 (19%) of these men went on to have surgery during the 18 months' follow-up.
d Patient agreement missing for one man in the routine care arm.
e 2/6 (33%) of these men went on to have surgery during the 18 months' follow-up.

surgery, 405 men (94%) did not have surgery, 21 men (5%) did go on to have surgery and for 5 men (1%), we do not know if they received surgery (because of withdrawal/death/unable to access medical notes).

Key secondary result: surgery rates

The proportion of men receiving surgery, regardless of surgery recommendation or listing, can be found in *Table 11*. The key secondary hypothesis for this trial was that surgery rates would be reduced in the UDS arm. The results show that there was no difference in the surgery rate at 18 months [153/408 (38%) men in the UDS arm had surgery, compared with 138/384 (36%) in the routine care arm]. This was well below the estimated surgery rate used in the sample size calculation (60–73%). We know the recommendation and surgery outcome for 748 men in total, and 26% of men who were recommended to have surgery did not receive it. Of those men not initially recommended for surgery, 4% did go on to receive it (e.g. after additional review due to prolonged symptoms). Of those men listed for surgery, 34 men (11%) did not undergo surgery during the 18-month follow-up; at least 10 of these 34 participants were still awaiting their surgery at their 18-month follow-up appointment (see *Time to surgery*).

Time to surgery

There were 291 men who had surgery for LUTS during UPSTREAM. In the UDS arm, the median time from randomisation to surgery was 216 days (7 months). In the routine care arm, the median time from randomisation to surgery was 177 days (6 months).

We are aware of at least 19 cases in which a surgical procedure for LUTS had been planned but was not carried out within the 18-month follow-up period. As mentioned above, 10 of these cases were listed for surgery during their decision-making appointment. Common reasons include medication failure, men having commitments and/or wanting more time to consider the options, health issues (e.g. high glucose levels) that have led to a justifiable delay in surgery, and slow pathway progress.

Types of surgery and corresponding International Prostate Symptom Score

Given the pragmatic nature of UPSTREAM, no operations were specified for use to treat voiding LUTS; therefore, there was a large variation of treatments (*Figure 10*). The majority were for bipolar or monopolar TURP.

TABLE 11 Secondary outcome: proportion of men having surgery, difference between arms

Variable	UDS, *n* (%)	Routine care, *n* (%)	OR[a] (95% CI)	*p*-value[a]
Surgery outcome				
Surgery conducted	153 (38)	138 (36)	1.05 (0.77 to 1.43)	0.741
No surgery	255 (63)	246 (64)		
Surgery outcome (if matching the surgeon's recommendation)[b]				
Surgery conducted	143 (44)	132 (41)	1.10 (0.79 to 1.54)	0.578
No surgery	185 (56)	189 (59)		
Recommended surgery				
Surgery conducted	143 (75)	132 (73)		
No surgery	47 (25)	48 (27)		
Not recommended surgery				
Surgery conducted	10 (5)	6 (3)		
No surgery	185 (95)	189 (97)		

OR, odds ratio.
a Adjusted for centre.
b Only those men who followed their surgeon's advice were included in this secondary analysis (e.g. if the surgeon recommended surgery and the man had it). One centre was dropped from the model due to perfect prediction.
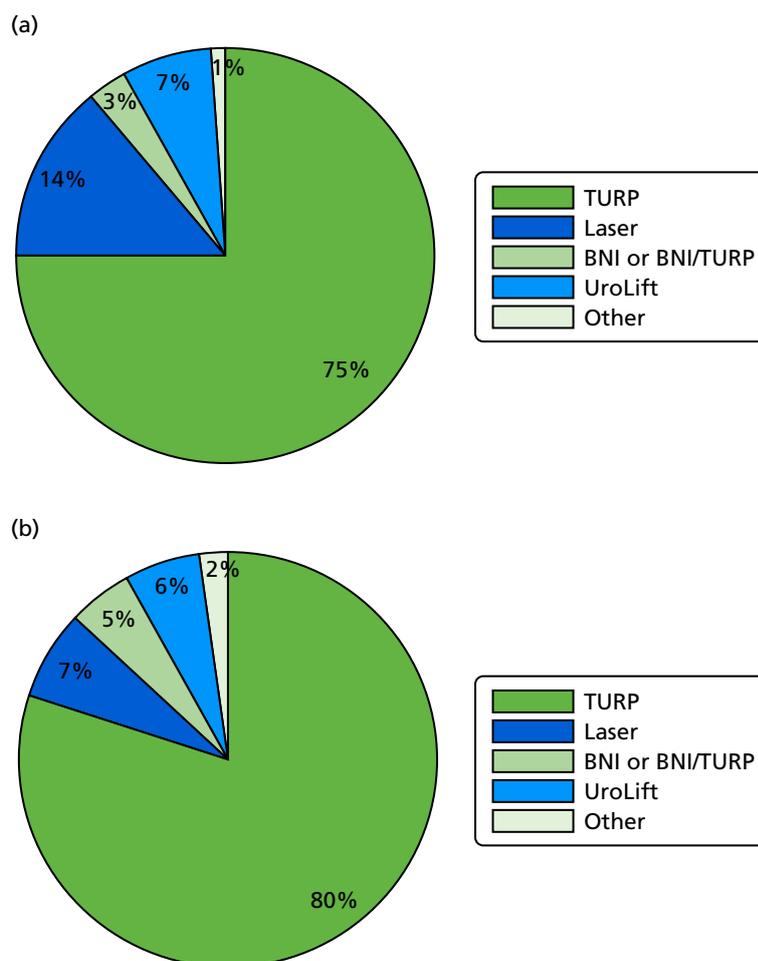
FIGURE 10 Types of surgery conducted in both arms of UPSTREAM. (a) UDS; and (b) routine care.

The most clear-cut improvement in LUTS was seen in those men who underwent surgery for their LUTS as recommended. As the primary research hypothesis was based around avoiding unnecessary LUTS surgery, we looked at the impact of surgery by treatment arm (*Table 12*), as well as the interaction between them (*Table 13*). The 'other' category has been excluded, as only one man had both a baseline and 18-month IPSS.

The improvement in IPSS seen in those men receiving surgery overall was 11.60, starting with an initial mean IPSS of 20.72 and ending with an 18-month mean IPSS of 9.11 (*n* = 253). The improvement in the routine care arm was 13.21 (*n* = 116) and in the UDS arm it was 10.25 (*n* = 116). For those men not receiving LUTS surgery, the drop in IPSS for the UDS arm was 3.09 (*n* = 191), and for the routine care arm it was 2.09 (*n* = 197). When then looking at the interaction between surgery and intervention arm on IPSS, there did appear to be an interaction (see *Table 13*). However, when using a per-protocol analysis, this effect was weakened slightly, suggesting that this exploration may have produced a chance finding.

A flow rate was taken approximately 4 months (± 1 month) after surgery, to assess the quality of the surgery. These data are explored by arm and surgery type in *Appendix 4*.

TABLE 12 Surgery types and IPSSs

Group	UDS			Routine care		
	n (%)	Baseline, mean (SD)	18 months, mean (SD)	*n* (%)	Baseline, mean (SD)	18 months, mean (SD)
Did not have surgery for LUTS						
Did not have LUTS surgery (all)	191	17.68 (6.57)	14.58 (7.18)	197	17.51 (7.25)	15.42 (7.60)
Recommended for surgery	37	18.95 (6.01)	15.49 (6.45)	38	19.47 (7.30)	17.76 (8.03)
Not recommended for surgery	151	17.26 (6.66)	14.27 (7.35)	155	16.90 (7.11)	14.66 (7.28)
Had surgery for LUTS						
Had surgery for LUTS (all)	137	19.64 (7.26)	9.39 (7.68)	116	22.00 (6.12)	8.79 (6.26)
Recommended for surgery	129	19.80 (7.34)	9.17 (7.53)	111	22.34 (5.86)	8.81 (6.34)
Not recommended for surgery	8	17.00 (5.48)	12.88 (9.72)	5	14.40 (7.57)	8.40 (4.45)
TURP (any type)	102	19.64 (7.01)	8.99 (7.62)	99	21.91 (6.31)	8.59 (5.96)
Laser (any type)	19	19.16 (8.30)	8.21 (7.61)	5	23.40 (7.40)	8.40 (3.51)
BNI	5	20.00 (12.14)	12.6 (8.26)	4	21.50 (5.80)	6.50 (7.77)
UroLift	10	20.40 (6.26)	13.2 (7.61)	7	22.43 (3.51)	10.29 (7.06)

TABLE 13 Interaction between surgery and treatment pathway

Variable	*n* (surgery : no surgery)	IPSS at 18 months	Interaction effect	
		Subgroup-specific difference, mean (95% CI)[a]	Difference, mean (95% CI)[a]	*p*-value
Subgroup analyses				
Arm of the trial				
UDS	137 : 191	−6.27 (−7.83 to −4.71)	2.93 (0.86 to 5.01)	0.006
Routine care	116 : 197	−8.91 (−10.61 to −7.22)		
As treated				
UDS	136 : 186	−7.26 (−8.84 to −5.69)	1.14 (−0.94 to 3.22)	0.283
Routine care	117 : 202	−7.70 (−9.39 to −6.01)		

a Adjusted for centre and baseline IPSS.

Conservative treatments and new medications over the 18 months

As well as surgery for LUTS, details of conservative treatments and new medications (since the initial treatment decision) were also collected at 18 months. Overall, 175 men received at least one conservative treatment (e.g. fluid advice). Twenty-one per cent of men ($n = 85/408$) in the UDS arm received a conservative treatment, compared with 23% in the routine care arm. Fewer men had conservative treatments when they had also had surgery (18% vs. 24% for men who did and men who did not have surgery, respectively). This was also true for new medications received over the 18-month period (e.g. tamsulosin and finasteride). Overall, 215 men received at least one new medication [115 men (28%) in the UDS arm and 100 men (26%) in the routine care arm]. Fewer men received new medications when they had also received surgery (20% vs. 31% for men who did and men who did not receive surgery, respectively).

Secondary results: adverse events

Adverse event data were collected for all men in the trial, excluding those who withdrew all their data ($n = 424$ and $n = 389$ for UDS and routine care arms, respectively). There were a total of 428 AEs in UPSTREAM (234 in the UDS arm and 194 in the routine care arm). The relationship to diagnostic testing and/or treatment was assessed along with the expectedness of the event, severity of event and Clavien–Dindo grading for surgery-related events. Each of these factors was then reviewed by a clinician independent of the trial, to ensure consistency of reporting across centres and to check for bias in classification. During the review, the Clavien–Dindo grading of two surgery-related events was queried, as the reviewer felt that they had been graded higher than necessary. After obtaining further details from the centres, these were re-graded. The proportion of events related to diagnostic tests or treatment/surgery was very similar between the arms, at 43–44% (*Table 14*).

Of the 241 unrelated AEs, 124 (51%) were serious. There were more serious and unrelated AEs in the UDS arm than in the routine care arm, 76 (six expected) and 48 (four expected) respectively. Of the remaining 117 non-serious and unrelated AEs, 54 (nine expected) were in the UDS arm and 63 (17 expected) were in the routine care arm.

When looking at the number of related events on both an ITT basis and as-treated basis, we can see that the number of related SAEs was very similar across the two arms (*Table 15*). Differences in the number of AEs between the arms were difficult to test at the event level. In the routine care arm, for example, one individual man suffered 7 of the 45 surgery-related and expected AEs. Formal tests were not carried out on these data but were, instead, utilised in testing if the number of events per man differed between the arms (*Table 16*). When looking specifically at acute urinary retention cases, there appeared to be a higher number in the routine care arm.

When looking at the number of events per man, there was no evidence to suggest that having UDS led to more or less AEs events than routine care.

TABLE 14 Number of total (serious and non-serious) AEs: relationship to a procedure

Variable	UDS, *n* (%)	Routine care, *n* (%)	*p*-value[a]
Was the event related to diagnostic tests or treatment/surgery?			
Probably	82 (35)	71 (37)	
Possibly	22 (9)	12 (6)	0.621
Unrelated	130 (56)	111 (57)	

a Ordinal logistic regression, adjusted for centre.

TABLE 15 Related AE details

Variable	UDS (ITT), number of events	Routine care (ITT), number of events	UDS (AT), number of events	Routine care (AT), number of events
Number of men with data[a]	424[b]	389[c]	381[d]	428[e]
SAEs[f]				
Due to UDS	3	1	3	1[g]
Due to surgery	21	23	20	24
Due to diagnostic tests	1	3	1	3
Due to other treatment[g]	2	0	2	0
Expected (non-serious) AEs				
Due to UDS	10	1	10	1[g]
Due to surgery	57	45	55	47
Due to diagnostic tests	0	0	0	0
Due to other treatment[g]	4	3	4	3
Unexpected (non-serious) AEs				
Due to UDS	3	0	1	2[h]
Due to surgery	3	4	3	4
Due to diagnostic tests	0	0	0	0
Due to other treatment[g]	0	3	0	3
Clavien–Dindo scores for surgery-related events (SAEs and AEs)				
Grade I	39	26	38	27
Grade II	35	33	35	33
Grade III	4	5	3	6
Grade IIIa	1	1	0	2
Grade IIIb	2	7	2	7
Grade IV	0	0	0	0
Grade IVa	0	0	0	0
Grade IVb	0	0	0	0
Grade V	0	0	0	0
Was the SAE/AE an episode of acute urinary retention?				
Acute urinary retention	13	29	11	31

a Excludes men who withdrew all data.
b Includes eight men who declined notes review, therefore not complete follow-up period.
c Includes three men who declined notes review, therefore not complete follow-up period.
d Includes four men who declined notes review.
e Includes seven men who declined notes review.
f All related SAEs were expected apart from one unexpected surgery-related event in the UDS arm (renal colic and hyponatraemia).
g Other treatments include ISC, catheter, flow test and medication.
h UDS procedure started but aborted (e.g. unable to catheterise, patient nervous).

Secondary results: patient-reported outcome measures

Overall PROMs were similar for the two groups, with fewer men reporting symptoms at 18 months (compared with baseline) for each of the urinary items. For nocturia (getting up to urinate more than once per night), there was moderate evidence to suggest that those men randomised to UDS fared better than those randomised to routine care (*Tables 17* and *18*). However, given the large number of secondary analyses, this may be a chance finding.

TABLE 16 Adverse events per man (SAEs and AEs)

Variable	UDS, *n* (%)	Routine care, *n* (%)	OR[a] (95% CI)	*p*-value[a]
All AEs				
No AEs	290 (68)	273 (70)		
One AE	77 (18)	72 (19)	1.08 (0.79 to 1.48)	0.609
More than one AE	57 (13)	44 (11)		
SAEs				
No SAEs	352 (83)	325 (84)		
One SAE	51 (12)	55 (14)	1.08 (0.74 to 1.59)	0.679
More than one SAE	21 (5)	9 (2)		
AEs				
No AE	333 (79)	316 (81)		
One AE	67 (16)	48 (12)	1.11 (0.77 to 1.61)	0.577
More than one AE	24 (6)	25 (6)		

a Adjusted for centre.

TABLE 17 Secondary outcome: ICIQ-MLUTS scores at baseline and at 18 months

Variable	Time of measurement[a]	*N* (UDS : routine care)	UDS, *n* (%) or mean [SD]	Routine care, *n* (%) or mean [SD]
ICIQ-MLUTS				
Voiding subscale[b]	Baseline	285 : 268	8.93 [4.03]	9.13 [4.18]
Voiding subscale[b]	18 months	296 : 278	6.41 [4.40]	6.19 [4.23]
Incontinence subscale[c]	Baseline	284 : 272	4.89 [3.22]	5.06 [3.07]
Incontinence subscale[c]	18 months	295 : 282	3.87 [3.07]	4.04 [2.81]
ICIQ-MLUTS: bother scores				
Daytime frequency (more than eight times)	Baseline	287 : 276	107 (37)	114 (41)
Daytime frequency (more than eight times)	18 months	297 : 284	84 (28)	75 (26)
Nocturia (more than one time per night)	Baseline	288 : 274	215 (75)	214 (78)
Nocturia (more than one time per night)	18 months	299 : 282	176 (59)	189 (67)

a Baseline measures were used only when the 18-month data were also available, to allow a fairer comparison.
b On a scale of 0–20, with larger scores indicating more severe symptoms.
c On a scale of 0–24, with larger scores indicating more severe symptoms.

Sexual symptoms were also similar between the arms. There were more men suffering with a reduced ejaculation by the end of the 18 months, with a greater increase seen in the UDS arm. There was also a small increase in the number of men reporting that urinary symptoms had affected their sex life (*Table 19*). However, when looking at the number of men suffering with sexual dysfunction at 18 months, there was no evidence of a difference; $p \geq 0.229$ (*Table 20*).

Secondary results: maximum urinary flow rate
Maximum urinary flow rate improved over the 18-month period in both arms of the trial. At 18 months, there was no difference seen in Q_{max} between the arms (*Table 21*).

TABLE 18 Secondary outcome: ICIQ-MLUTS analysis

Variable	N (UDS : routine care)	UDS, n (%) or mean [SD]	Routine care, n (%) or mean [SD]	Difference in means[a] (95% CI)	p-value[a]
ICIQ-MLUTS					
Voiding subscale[b]	296 : 278	6.41 [4.40]	6.19 [4.23]	0.09 (–0.59 to 0.77)	0.791
Incontinence subscale[c]	295 : 282	3.87 [3.07]	4.04 [2.81]	–0.27 (–0.67 to 0.13)	0.191
ICIQ-MLUTS: bother scores					
Daytime frequency (more than eight times)	297 : 284	84 (28)	75 (26)	1.12 (0.75 to 1.67)	0.573
Nocturia (more than one time per night)	299 : 282	176 (59)	189 (67)	0.66 (0.44 to 0.99)	0.042

a Adjusted for centre and baseline scores; centre was not adjusted for in the binary outcomes, due to perfect prediction.
b On a scale of 0–20, with larger scores indicating more severe symptoms.
c On a scale of 0–24, with larger scores indicating more severe symptoms.

TABLE 19 Secondary outcome: ICIQ-MLUTSsex scores at baseline and at 18 months

Variable	Time of measurement[a]	N (UDS : routine care)	UDS, n (%)	Routine care, n (%)
ICIQ-MLUTSsex				
Erections (reduced or none)	Baseline	271 : 261	190 (70)	193 (74)
Erections (reduced or none)	18 months	287 : 270	206 (72)	196 (73)
Ejaculation (reduced or none)	Baseline	265 : 253	206 (78)	209 (83)
Ejaculation (reduced or none)	18 months	286 : 264	244 (85)	219 (83)
Painful ejaculation (yes)	Baseline	233 : 227	38 (16)	49 (22)
Painful ejaculation (yes)	18 months	255 : 246	43 (17)	39 (16)
Urinary symptoms affected sex life?	Baseline	258 : 252	176 (68)	167 (66)
Urinary symptoms affected sex life?	18 months	274 : 266	197 (72)	179 (67)

a Baseline measures were used only when the 18-month data were also taken, to allow a fairer comparison.

TABLE 20 Secondary outcome: ICIQ-MLUTSsex analysis

Variable	N (UDS : routine care)	UDS, n (%)	Routine care, n (%)	OR[a] (95% CI)	p-value[a]
ICIQ-MLUTSsex					
Erections (reduced or none)	287 : 270	206 (72)	196 (73)	1.03 (0.63 to 1.68)	0.898
Ejaculation (reduced or none)	286 : 264	244 (85)	219 (83)	1.37 (0.82 to 2.28)	0.229
Painful ejaculation (yes)	255 : 246	43 (17)	39 (16)	1.16 (0.67 to 1.98)	0.598
Urinary symptoms affected sex life?	274 : 266	197 (72)	179 (67)	1.28 (0.84 to 1.96)	0.252

OR, odds ratio.
a Adjusted for baseline scores but not centre, due to perfect prediction.

TABLE 21 Secondary outcome: Q_{max} at 18 months, difference between arms

Variable	n (UDS : routine care)	UDS, mean (SD)	Routine care, mean (SD)	Difference in means[a] (95% CI)	p-value[a]
Q_{max} score					
Q_{max} at 18 months	268 : 270	15.26 (8.25)	15.68 (9.10)	0.14 (–1.33 to 1.62)	0.850
Baseline Q_{max}	260 : 260	11.85 (6.28)	12.39 (6.83)		

a Adjusted for centre and baseline Q_{max}.

Secondary results: urodynamics satisfaction questionnaire

Overall, satisfaction with UDS was very good. The median satisfaction score was 10 out of 10. Fifty per cent of men felt that the test was better than expected; however, 26% reported that it was worse than expected. The satisfaction with each element of the care given was very high. The lowest proportion of satisfaction was found in the explanation of the results, with 11% saying that they were not satisfied (*Table 22*).

TABLE 22 Secondary outcome: satisfaction with UDS

Satisfaction question	N (UDS : routine care)	UDS arm, n (%) or median (IQR)	Routine care arm, n (%) or median (IQR)
Number of men who received UDS		353	28
ICIQ-UDS-S			
Overall satisfaction (0–10)	297 : 6	10.00 (9.00–10.00)	10.00 (10.00–10.00)
The test was better than expected	302 : 6	152 (50)	3 (50)
The test was same as expected		52 (17)	1 (17)
The test was worse than expected		77 (26)	2 (33)
The test was different but no better or worse		21 (7)	0 (0)
Did you think the test was successful?[a]	282 : 5	275 (98)	5 (100)
Knowing what you know now, would you take the test?[b]	303 : 6	294 (97)	6 (100)
Satisfaction with information received in the post?[c]	293 : 6	260 (89)	6 (100)
Satisfaction with information from the doctor?[c]	284 : 6	268 (94)	6 (100)
Satisfaction with the doctor?[c]	277 : 5	274 (99)	5 (100)
Satisfaction with the nurse who performed the test?[c]	297 : 6	294 (99)	6 (100)
Was your privacy and dignity preserved?[d]	298 : 6	294 (99)	6 (100)
Satisfaction with the explanation of the results?[e]	293 : 5	261 (89)	5 (100)
Would you recommend the test to friends/family?[f]	301 : 6	278 (92)	6 (100)

IQR, interquartile range.
a Very, quite, somewhat or a little successful versus unsuccessful.
b Definitely or probably versus not sure, probably or definitely not.
c Very-little satisfied versus neutral or dissatisfied.
d Extremely or moderately preserved versus a little bit/not at all.
e Very-little satisfied versus neutral, dissatisfied or not received.
f Definitely or probably versus not sure or not.
Reprinted from *European Urology Focus*, vol. 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). pp. 340–50, copyright 2019,[27] with permission from Elsevier.

Statistical ancillary analyses

Subgroup analyses

Prespecified formal tests of interaction were employed to explore potential effect modifiers. There was no evidence to suggest that there were any interactions between any of the subgroups on the effect of trial arm on IPSS (*Tables 23* and *24*).

Sensitivity analyses

Several sensitivity analyses were conducted to test the robustness of the primary outcome results. Two of these analyses were conducted to check that the timing of surgery did not influence 18-month IPSSs. The median time from randomisation to LUTS surgery in the UDS arm was 215 days (7 months) and the median time to LUTS surgery in the routine care arm was 181 days (6 months). The median time from surgery to completing the 18-month questionnaire was 364 days (12 months) in the UDS arm and 407 days (13 months) in the routine care arm. Adjusting for time since surgery led to CIs that went beyond the non-inferiority margin, giving results that were inconclusive with regards to non-inferiority.

During the peer-review process, when submitting the SAP for publication, the reviewers requested a repeated-measures analysis. The IPSSs were lower for the UDS arm at all time points, except for 6 months when the routine care arm appeared to have the lowest IPSSs (*Figure 11*). As surgery had such a positive impact on IPSSs, and men in the routine care arm were having surgery slightly earlier, this may have caused the sharper decline in IPSSs at 6 months. The results from the repeated-measures model still showed non-inferiority at 12 months and at 18 months (*Table 25*).

TABLE 23 Subgroup descriptives

Variable	n (UDS : routine care)[a]	UDS, mean (SD)		Routine care, mean (SD)	
		IPSS at baseline	IPSS at 18 months	IPSS at baseline	IPSS at 18 months
Subgroup descriptives					
Age (years)					
≤ Median	167 : 157	18.70 (7.25)	12.92 (8.36)	19.83 (7.11)	13.73 (8.08)
> Median	161 : 156	18.28 (6.58)	11.89 (7.19)	18.51 (7.21)	12.20 (7.48)
Flow rate (ml/second)					
≤ 12	200 : 183	19.19 (6.74)	12.05 (7.85)	19.94 (7.23)	12.56 (8.22)
> 12	118 : 118	17.22 (7.14)	13.13 (7.77)	17.69 (7.03)	13.14 (6.69)
Maximum VV (ml)					
< 200	138 : 134	19.21 (7.02)	12.48 (7.65)	19.62 (7.29)	12.67 (8.12)
≥ 200	182 : 172	17.92 (6.83)	12.37 (7.95)	18.78 (7.03)	13.08 (7.49)
Storage dysfunction					
No nocturia	77 : 65	15.34 (5.73)	10.65 (6.71)	14.82 (6.90)	10.35 (6.89)
Nocturia	241 : 243	19.35 (6.99)	12.83 (8.04)	20.33 (6.82)	13.70 (7.96)
Severity of storage LUTS[b]					
Less substantial	189 : 174	14.67 (5.39)	11.14 (7.29)	15.02 (5.99)	11.32 (6.76)
More substantial	139 : 139	23.69 (5.18)	14.14 (8.19)	24.37 (4.76)	15.02 (8.54)

a The numbers refer to the number of men with IPSS scores at baseline and 18 months.
b The summation of items 2, 4 and 7 in the IPSS questionnaire (split by the median).

TABLE 24 Subgroup analyses

Variable	n (UDS : routine care)[a]	IPSS at 18 months[b]	Interaction effect[b]	
		Subgroup-specific difference in means (95% CI)	Difference in means (95% CI)	p-value
Subgroup analyses *Age (years)*				
≤ Median	167 : 157	–0.16 (–1.88 to 1.56)		
> Median	161 : 156	–0.47 (–2.00 to 1.06)	–0.33 (–2.60 to 1.94)	0.773
Flow rate (ml/second)				
≤ 12	200 : 183	–0.54 (–2.06 to 0.98)		
> 12	118 : 118	0.21 (–1.65 to 2.08)	0.54 (–1.84 to 2.92)	0.649
Maximum VV (ml)				
< 200	138 : 134	–0.61 (–2.43 to 1.20)		
≥ 200	182 : 172	–0.41 (–1.97 to 1.15)	0.35 (–1.99 to 2.69)	0.763
Storage dysfunction				
No nocturia	77 : 65	–0.30 (–2.56 to 1.97)		
Nocturia	241 : 243	0.49 (–1.85 to 0.88)	–0.60 (–3.33 to 2.14)	0.661
Severity of storage LUTS[c]				
Less substantial	189 : 174	–0.14 (–1.49 to 1.20)		
More substantial	139 : 139	–0.61 (–2.63 to 1.42)	–0.70 (–2.99 to 1.60)	0.542

a The numbers refer to the number of men with IPSS scores at baseline and 18 months.
b Coefficient for the interaction between arm and subgroup, along with a likelihood ratio test *p*-value, comparing models including/excluding the interaction term, adjusting for centre and baseline diagnosis.
c The summation of items 2, 4 and 7 in the IPSS questionnaire (split by the median).
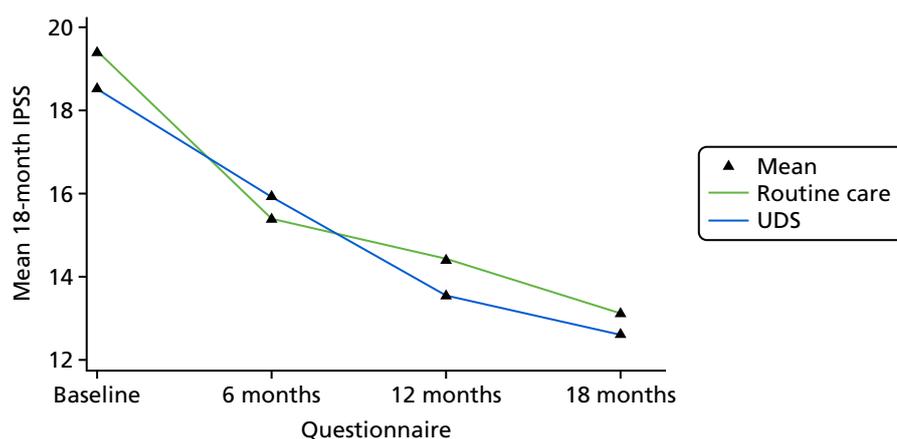
FIGURE 11 International Prostate Symptom Scores over time in UPSTREAM.

The per-protocol and CACE analyses (although prone to bias), were similar to the ITT results. The per-protocol analysis upper CI was 1.05, slightly above the non-inferiority margin, but this could be attributed to the reduction in sample size. As no baseline variables differed by 10% or 0.5 SDs, or more, the sensitivity analysis that planned to adjust for imbalance at baseline was not carried out. Adjustment for prespecified confounders gave very similar results to the primary analysis (see *Table 25*).

TABLE 25 Sensitivity analyses: IPSSs, difference between arms

Variable	n (U : R)[a]	UDS, mean (SD)	Routine care, mean (SD)	Difference in means[b] (95% CI)
Primary analysis				
Intention to treat[c]	340 : 329	12.61 (7.92)	13.11 (7.86)	–0.33 (–1.47 to 0.80)
Sensitivity analyses				
Per protocol[d]	306 : 305	12.56 (7.80)	12.78 (7.50)	–0.10 (–1.25 to 1.05)
CACE analysis[e]	–	–	–	–0.39 (–1.70 to 0.92)
Repeated measures[f]				
At 6 months	310 : 272	15.91 (7.44)	15.39 (7.43)	0.92 (–0.23 to 2.06)
At 12 months	300 : 263	13.54 (7.84)	14.40 (7.89)	–0.64 (–1.80 to 0.51)
At 18 months	340 : 329	12.61 (7.92)	13.11 (7.86)	–0.19 (–1.28 to 0.90)
LOCF using 6 and 12[g]	346 : 339	12.55 (7.96)	13.25 (7.98)	–0.63 (–1.76 to 0.50)
Multiple imputation[h]	427 : 393	12.80 (8.44)	13.19 (8.36)	–0.11 (–1.22 to 0.99)
Adjusted for potential confounders[i]	326 : 307	–	–	–0.30 (–1.44 to 0.84)
Adjusted for time since surgery[j]	336 : 323	–	–	0.19 (–0.84 to 1.21)
Post hoc: removal of surgery[k]	325 : 316	12.70 (7.94)	13.13 (7.92)	–0.24 (–1.39 to 0.92)
Post hoc: follow-up window[l]	278 : 250	12.51 (7.90)	12.87 (7.63)	–0.05 (–1.29 to 1.19)

LOCF, Last Observation Carried Forward; R, Routine care; U, Urodynamics.
a The number of men with available 18-month IPSS data.
b Adjusted for centre and baseline IPSS.
c 12 : 16 men, in the urodynamics and routine care arms, 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 (7 : 15 excluded from analysis due to missing baseline IPSS data).
e Unbiased estimates to account for patient crossover using instrumental–variables regression (12 : 16 excluded from analysis due to missing baseline IPSS data).
f An interaction term was included in the mixed-effects model, therefore three coefficients have been presented, one for each time point. Excluding the interaction gave an overall coefficient of 0.02 (95% CI –0.87 to 0.91), $n = 362 : 335$.
g Last Observation Carried Forward: where missing baseline data is imputed using 6-month questionnaire data and 18-month data is imputed using 12-month questionnaire data (13 : 7 excluded from analysis due to missing baseline/6-month IPSS data).
h Multiple imputation by chained equations.
i Additional adjustment for age and presence of comorbidities (14 : 22 excluded from analysis due to missing baseline covariate data).
j Adjusted for the number of days between the surgical procedure and the 18-month questionnaire, a value of 1000 days was imputed where no surgery took place and men were excluded if we didn't know whether or not they had received surgery (12 : 16 excluded from analysis due to missing baseline data).
k As a post hoc analysis, the team excluded any men that had surgery within 3 months of completing their IPSS questionnaire (12 : 16 excluded from analysis due to missing baseline IPSS data).
l Men who completed their 18-month follow-up outside a 16- to 20-month post-randomisation window and men who completed their baseline questionnaire outside a six-month pre- and post-randomisation window were excluded (4 : 6 excluded from analysis due to missing baseline IPSS data).

Adjustment for the time between surgery and completion of the questionnaire led to results that were inconclusive. When men had not received surgery, the value of 1000 days was imputed, and, when we did not know if a man had received surgery, their data were excluded. The use of 1000 days was arbitrary and was used to ensure that anyone not receiving surgery was still included in the analysis, but had values far enough away from those men who received surgery. These are not robust assumptions; therefore, these results should be viewed with caution. Removing surgery within 3 months of the questionnaire was far more robust, as all men included had at least a 3-month post-surgery gap, to allow any short-term surgery effects to subside. Restricting the follow-up window to 16–20 months also provided results that were inconclusive, as the CI stretched over the non-inferiority margin. Imputation using 6- and 12-month IPSS

allowed an additional six men to be analysed in the UDS arm and an additional 10 men in the routine care arm. Results from this analysis also concluded non-inferiority. However, the CI from the multiple imputation analysis had an upper limit of 0.99, just within the non-inferiority margin.

Multiple imputation details

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 arm of the trial (binary); whether or not UDS was received (binary); whether or not men received surgery (binary); age (continuous); comorbidities (binary); centre (categorical); Q_{max} at baseline and at 6, 12 and 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 was used for the imputation.

Statistical results summary

- Primary analysis results showed that UDS was non-inferior to routine care for IPSSs at 18 months of follow-up, with a CI within the non-inferiority margin of 1 point.
- Overall, for both arms, IPSS dropped from a mean of 18.94 ($n = 774$) to a mean of 12.86 ($n = 669$).
- The anticipated advantage of lower surgery rates in the UDS arm was not found. Surgery rates were lower than expected in the routine care arm and similar between the arms, 38% and 36% for the UDS and routine care arms, respectively. In an exploratory analysis, we found, in both arms, that having surgery led to the greatest reduction in IPSSs at 18 months, compared with not having surgery.
- The number of AEs experienced per person was similar between the arms. Unrelated AEs appeared to be more numerous in the UDS arm. Thirty-two per cent of men in the UDS arm suffered at least one AE, compared with 30% in the routine care arm.
- As with the IPSSs, PROMs for urinary symptoms dropped in both arms. There was evidence to suggest that men in the UDS arm experienced a greater reduction in nocturia than men in the routine care arm ($p = 0.010$). However, given the large number of secondary analyses carried out and lack of additional urinary symptom benefits, this may be a chance finding.
- PROMs for sexual symptoms showed no obvious differences between the arms at 18 months, but there did appear to be an increase in the number of men suffering from reduced ejaculation in the UDS arm. However, the CI was wide enough to include an odds ratio of 1 and, given the large number of tests carried out, this may be a chance finding.
- Satisfaction with UDS was high in all men who received it, with 98% agreeing that the test was successful and 97% saying that they would have the test again.
- There was no evidence to suggest that subgroup interactions were present, although the study was underpowered to detect these.
- The sensitivity analyses carried out agreed with the primary analysis. However, there were at least two occasions when the CI included the non-inferiority margin, rendering an inconclusive result. Given the large number of sensitivity analyses carried out, this is perhaps unsurprising.

Chapter 4 Economic evaluation

Introduction

The objective of this section is to report the within-trial economic evaluation of UPSTREAM, conducted from randomisation to 18-month follow-up, from three perspectives: (1) the NHS secondary care perspective (defined as data on secondary care obtained from local routinely collected hospital data), (2) a wider NHS perspective and (3) a patient perspective. Cost-effectiveness analyses of the assessment of patients with LUTS following a UDS pathway compared with the routine care pathway, from the three perspectives in relation to quality-adjusted life-years (QALYs), are reported. The time horizon for the cost-effectiveness analysis was 18 months, the duration of the follow-up period. Costs are derived from resources used by men in relation to secondary care and community-based NHS services and any out-of-pocket expenditure they may have experienced in the course of the treatment of their LUTS. QALYs are determined from the EuroQol-5 Dimensions, five-level version (EQ-5D-5L), questionnaire and the cross-walk valuation set.[31]

Methods

Measurement and valuation of relevant resource use
Resource use was collected from randomisation for 18 months. Data on resource use for this analysis came from two main sources: Patient Level Information and Costing System and patient-completed questionnaires.

Use of routine hospital data
Informatics or similar departments in all centres were approached and independent information analysts were asked to provide electronic information from routinely collected hospital data. Routine hospital data were obtained for those men who had granted ongoing consent for the trial team to access their medical notes. We requested diagnostic codes (*International Statistical Classification of Diseases*, Tenth Edition), procedural codes (International Classification of Interventions and Procedures), currency codes [Healthcare Resource Group (HRG)], and admission and discharge dates at finished consultant episode (FCE)-level data, for all for inpatient stays (including day cases). For outpatient attendances and outpatient diagnostic imaging tests, we requested, as applicable, procedure, HRG/currency codes, diagnostic codes and attendance dates (although we requested diagnostic codes, these were widely unavailable). We requested attendance date and HRG code data for emergency admissions.

For all centres, a Microsoft Excel® spreadsheet (Microsoft Corporation, Redmond, WA, USA), containing a list of trial (participant) identifiers and individualised dates of randomisation and 18-month follow-up, was sent to the onsite trial research nurses; they then added date of birth and NHS and hospital numbers to the trial identifiers, and securely forwarded the information to the information analyst. Data on all inpatient stays, outpatient and emergency department attendances, and diagnostic imaging outpatient appointments between the two given dates were added to this spreadsheet, which was then returned to the research nurse. All data were anonymised by the research nurses prior to the secure transfer of data to the trial team.

For 58 participants completing follow-up close to the trial end date, it was necessary to request their data before their 18-month follow-up date. We accepted a minimum of 15 months' follow-up data for these patients, effectively administratively censoring these patients at 15 months.

Two clinicians devised a list of all specialties that may conceivably be involved in treating conditions that may present with LUTS (*Box 3*). All events recorded against specialties that, according to clinician opinion, would be unlikely to be related to the treatment of LUTS were deleted.

BOX 3 Specialties treating patients presenting with LUTS

- General surgery.[a]
- Anaesthetics.
- Endocrinology.
- Diabetology.
- Nephrology.
- Neurology (including neurophysiology, excluding stroke medicine).
- Neurophysiology.
- Medical oncology.
- Clinical oncology.
- Urology.

a General surgery FCEs were included in the inpatient data set only when a subsequent urology FCE was reported. Urology cases are often admitted under general surgery out of hours. General surgery outpatient attendances were not costed.

All inpatient data were processed using the 2016/17 NHS reference costs grouper.[32] This ensured that all HRGs generated from the diagnostic and procedure codes provided by hospitals could be subsequently valued using the published 2016/17 NHS Reference Costs.[33] The relevant diagnostic or procedure codes were used to assign a treatment specialty if this information was missing. If necessary, main specialty codes were derived from treating specialty codes (e.g. colorectal surgery treating specialty was assigned the general surgery main specialty). Urology is both a treating and a main specialty.

When the grouper generated HRG codes that matched with a zero cost in the reference costs, these were checked for errors and, when appropriate, were assigned an alternative HRG code based on clinical opinion. Unbundled diagnostic imaging, nuclear medicine and chemotherapy services produced by the grouper were costed separately and added to the secondary care costs.

Finished consultant episodes were amalgamated into spells. A spell consists of a continuous stay in a single NHS hospital trust. The NHS treats any two inpatient spells that occur in a short duration as a single spell. Any patients with spells that occurred within 48 hours of each other were costed as a single spell. If any FCE within a spell was recorded against a LUTS-related specialty, the spell was included in the inpatient resource use costing.

All day-case and inpatient admissions were assumed to be elective and costed as such, unless the patient had an emergency attendance recorded at the treating hospital in the 24 hours prior to an inpatient admission. If an inpatient stay was greater than the grouper-provided trim point, then an additional cost was applied to the number of excess bed-days. Any spells in which the patient was admitted and discharged on the same day were costed using day-case reference costs.

All outpatient attendances not assigned a HRG code were valued as consultant-led, follow-up, face-to-face appointments, unless the hospital indicated otherwise in writing. Any International Classification of Interventions and Procedures codes suggestive of an outpatient procedure appointment, such as UDS or uroflowmetry, rather than a clinic appointment, were recorded and valued as such. Outpatient procedures and clinic appointments occurring on the same day were costed separately.

Two centres were unable to provide HRG information of emergency department attendances. The date of attendance was provided and the most commonly used HRG from the rest of the data set was attributed to these attendances for costing purposes.

Resource use questionnaires

The resource use questionnaire was administered to the men at baseline and at 6, 12 and 18 months' follow-up. The questionnaire at the baseline and 18-month clinics was administered by the nurses if time permitted, otherwise it was given to the participants for them to complete in their own time, to return by post. At 6 and 12 months, the questionnaires were administered by post. The questionnaires were used to collect information on other hospital health-care use: NHS and private community-based health-care use [e.g. general practitioner (GP) visits, district nurse visits, community urology services] and medications, in addition to travel, time off work and any other expenses (e.g. incontinence pads, resulting from their condition and/or treatment).

All resource use and the unit costs used to value the resources are given in *Table 26*. 2016/17 costs excluding value added tax were used to value the resource use.

TABLE 26 Resources collected and their valuation

Resource	Unit cost (£)	Source of cost
Inpatient stays	HRG[a]	*NHS Reference Costs 2016/2017;*[33] Curtis and Burns[34]
Day cases	HRG[a]	*NHS Reference Costs 2016/2017*[33]
Outpatient visits	HRG[a]	*NHS Reference Costs 2016/2017*[33]
Outpatient procedures	HRG[a]	*NHS Reference Costs 2016/2017*[33]
Accident and emergency attendances	HRG[a]	*NHS Reference Costs 2016/2017*[33]
Outpatient diagnostic imaging	HRG[a]	*NHS Reference Costs 2016/2017*[33]
GP surgery visit	29.00[b,c]	Curtis and Burns[34]
GP home visit	89.44[b,c,d]	Curtis[35]
GP telephone call	14.60	Curtis and Burns[34]
GP nurse visit	5.53[c,e]	Curtis and Burns[34]
District nurse visit	38.68[d]	Curtis and Burns[36]
Community continence nurse visit	83.00	*NHS Reference Costs 2016/2017*[33]
Car mileage[f,j,k]	0.45	GOV.UK[37]
Other travel costs	Varies	Trial participant
Community-based urology service visit	103.00	*NHS Reference Costs 2016/2017*[33]
Prescription cost to NHS	7.80[g]	Curtis and Burns[34]
Prescription cost to patient	8.40[h]	NHS Commissioning Board[38]
Unpaid time off work	7.20[i]	Office for National Statistics[39]

a HRG codes were obtained directly from hospitals or derived using the NHS reference costs grouper.[32]
b Excluding direct care staff costs.
c Excluding qualification costs.
d The Hospital and Community Health Services index was used to inflate costs to 2016–17 prices.
e Based on the assumption of a 9.22-minute consultation.
f A participant who used a hospital or community health-care resource but did not provide details of a mode of transport were assumed to travel by car.
g Prescription costs in Curtis and Burns[34] are the estimated actual cost per medication net of the estimated rebate.
h Prescription cost up to April 2017: £8.40.
i Minimum wage, aged ≥ 25 years, October 2016–March 2017.
j Travel to GP surgery: based on a distance of 1.24 miles, from a 2015 national audit report that states that 92% of the population live within 2 km (1.24 miles) of a GP surgery.[40]
k Travel to hospital: based on an 8.8-mile-round trip, using the estimate from a study by Roberts *et al.*[41] that the average distance to an acute hospital with emergency care facilities is 7.2 km (4.4 miles).

Outcome data collection and valuation

As recommended by NICE,[42] the QALY is the primary economic outcome for UPSTREAM. The EQ-5D-5L was placed in the resource use questionnaire and was administered in the same way and at the same time points (baseline and 6, 12 and 18 months' follow-up). Men's self-reported EQ-5D-5L values at baseline and at 6, 12 and 18 months' follow-up were transformed into utility scores using the cross-walk valuation set, as currently recommended by NICE.[31] QALYs for each man are calculated from the utility scores using the area under the curve approach. The QALY calculation takes into account any deaths that have occurred during the duration of the trial. The primary outcome of the effectiveness analysis, IPSS, was evaluated in the cost–consequence analysis.

Data cleaning and missing data

The health economists were blinded to the randomisation allocation until all the data cleaning and valuation of resources had taken place for the NHS secondary care perspective.

Obtaining secondary care data from routinely collected hospital data was assumed to deliver a complete set of resource use (i.e. no resource was used if it was not reported). Therefore, any true missing data are due to an administrative error and not related to the trial intervention or the trial participant, and can be considered missing completely at random and unlikely to bias the estimate of cost-effectiveness.

In relation to the missing items from the resource use questionnaires, if the resource use questionnaire was returned and a resource use item was missing, it was assumed that no resource had been used. Completion of data across all three follow-up time points was poor for travel costs (other than travel by car) and for use of incontinence pads. In order for the multiple imputation to run, these items were removed from the analysis. For the remaining missing data, multiple imputation by chained equations was used to generate estimates of resource use. The model for multiple imputation included trial arm, age, centre, baseline Q_{max} and PVR measurements. In addition, the following variables were included at baseline and at 6, 12 and 18 months: IPSS; utility; primary and additional secondary care costs, and associated travel costs; secondary care costs (obtained from routine hospital data) for inpatient, outpatient, emergency department and diagnostic imaging for the duration of the trial, and associated travel costs; and NHS and personal costs for prescriptions and incontinence aids. Forty individual imputations were conducted and combined using Rubin's rules[43] in Stata for all analyses. The randomisation seed of 525 was employed to create reproducible imputations.

Analysis

The economic analyses are conducted under an ITT approach (i.e. analysing men in the arm to which they were randomised, irrespective of any post-randomisation changes). As the trial period fell beyond 1 year, both costs and effects falling in year 2 were discounted at 3.5%, in accordance with NICE guidance.[42] No modelling has been specified in this evaluation, as the work, when planned, was seen as a definitive trial, as it was expected that most uncertainty in relation to cost differences would be captured within the 18-month duration of the trial.

The cost of each item of resource used during the 18 months of follow-up is evaluated as the resource use (e.g. number of GP appointments) multiplied by its unit cost. The total cost for each resource category (e.g. outpatient attendances), the NHS secondary care, NHS and patient perspective for each individual man, was calculated as the sum of the cost of resource use items. The mean resource use and costs were estimated and presented by trial arm for each resource use category (e.g. outpatient attendances). The method of seemingly unrelated regressions (SURs), which accounts for the correlation between individual costs and QALYs, was used to estimate adjusted mean costs and QALYs and the differences in adjusted mean costs, QALYs (and their associated 95% CIs), between the trial arms in relation to secondary care NHS costs and NHS costs and patient costs. Costs and QALYs were adjusted for centre and baseline IPSS. In addition, costs were adjusted for baseline resource use and QALYs were adjusted for baseline utility.[44]

Cost–consequence analysis

A cost–consequence analysis was conducted in which adjusted NHS costs were compared with adjusted QALYs and IPSSs.

Cost-effectiveness analysis

Cost-effectiveness analyses were conducted in which the secondary care costs and the NHS costs were compared with QALYs. Incremental cost-effectiveness ratios (ICERs) were created using SUR if neither treatment was dominant (i.e. less expensive and more effective). SUR outputs were used to estimate the incremental net monetary benefit (INMB) statistic at the standard NICE willingness-to-pay threshold of £20,000 per QALY. Cost-effectiveness acceptability curves (CEACs), which show the probability that a pathway with UDS is the cost-effective option compared with routine care at different willingness to pay per QALY thresholds, were created to explore uncertainty.

Sensitivity analyses

A series of one-way sensitivity analyses were conducted primarily on the NHS secondary perspective analysis (as this contained the most complete data and the greatest cost drivers) to test the robustness of different parameter estimates and assumptions made in relation to resource use and costs.

The sensitivity analyses in relation to the secondary care perspective were as follows:

- Varying the discount rate between 0% and 5%.
- Secondary care resource use was re-estimated using the upper and lower quartile estimates of the reference costs, rather than the point estimate (calculated as the mean cost averaged across providers in England).
- Excluding participants who did not have a complete 18 months of follow-up.
- Censoring all participants' resource use and QALYs at 15 months of follow-up.
- Private provider costs were excluded from the analysis. In the base-case analysis they had been included as it was unclear whether the providers were contracted to provide NHS care through waiting list initiatives (therefore acting as a substitute for NHS care) or whether participants engaged private providers through direct payment or insurance (probably as a complement to NHS care).
- Additional costs for unidentified UDS tests from routine data.

To complete the analysis on time, it was necessary to administratively censor hospital resource use data collected for some men between 15 and 18 months. For 58 men, we accepted a minimum of 15-month follow-up data from hospital routine systems. This was done to ensure that the data were extracted and returned to the trial team with sufficient time to process the data, upload it to a separate health economic secure area within the REDCap[45] (Research Electronic Data Capture 8.10.19 © 2019 Vanderbilt University, TN, USA) database and conduct the analysis.

We present two sensitivity analyses investigating the impact of this administrative censoring on the analysis. In sensitivity analysis 3, we exclude the 58 patients for whom we did not have complete routine data follow-up. In sensitivity analysis 4, we censor all patients at 15-month follow-up. In this sensitivity analysis, we ignore any routine hospital-reported resource use between 15 and 18 months of follow-up for all patients. To calculate QALYs at 15 months, we use the utility values for 12 and 18 months and the area under the curve method, and weight them by 3 months instead of 6 months.

Information about diagnostic tests and relevant surgical procedures was collected from two sources: (1) routine hospital data and (2) research nurse-completed CRFs. When comparing these two data sets, we found that, for the UDS arm, fewer UDS tests were identified in the routine data than in the CRFs. This under-reporting may have arisen because some UDS tests were undertaken at another health-care facility. To investigate the effect of this underestimate, we performed a sensitivity analysis (sensitivity analysis 6), in which we applied a UDS cost to all those men who were reported to have UDS tests in the CRF and who did not have identifiable UDS tests recorded in the routine data.

Results

Numbers analysed

Resource use and cost data from routine hospital data, which formed the secondary care perspective, were obtained for 802 men for whom we had consent to access their medical records (98% of those men randomised). In the case of the NHS perspective, the completeness of the data declined to 15%. The number of complete cases from a patient perspective was 173 (22%). The EQ-5D-5L data were complete for 93% men at baseline, 72% at 6 months, 70% at 12 months and 77% at 18 months, which meant that complete QALY data were obtained for 454 (57%) participants. The SUR analyses were conducted on all randomised men for whom we had consent to use their data ($n = 802$).

Resource use and unadjusted costs

Participants in the UDS arm had more inpatient admissions and day-case spells than participants who received routine care (*Table 27*). There was little difference in the number of operations performed in each arm. From the routine data, we found that 30% of participants in the UDS pathway had surgery at the main treating hospital, compared with 31% in the routine care arm. Of note, twice as many participants in the UDS arm were treated with HoLEP compared with the routine care arm. Open prostatectomy with capsulectomy were performed only in participants in the UDS arm. Routine care participants were more likely to have day-case bipolar enucleation of the prostate.

Similar numbers of participants had repeated surgery during the 18-month follow up period (five participants in the routine care arm and four participants in the UDS arm). In the routine care arm, three participants had two monopolar or reaming enucleation (which we will call 'usual' TURPs), one participant had a usual TURP and a bipolar TURP and one participant had a usual TURP followed by a prostatectomy and capsulectomy. In the UDS arm, four participants had repeated surgery. Two participants had two usual TURPs each, one participant had a usual TURP and a BNI as separate operations and the final participant had a bipolar TURP and BNI as separate operations.

There was no difference in the number of participants presenting with urinary retention and requiring urethral catheterisation [7/416 (0.016%) of participants in the UDS arm, similar to the 7/386 (0.018%) participants in the routine care arm, one of whom was twice catheterised for urinary retention].

As expected, the number of outpatient attendances was greater in the UDS arm, reflecting the inclusion of UDS visits. There was no difference in the use of diagnostic imaging between the arms.

Resource use was similar between the two arms for additional secondary and primary care (*Table 28*).

In *Table 29*, showing the patient's perspective, the participants in the routine care arm have fewer medications prescribed but a higher cost. This is because one participant in this group did not have a pre-payment certificate; therefore, he paid more for his medications than the cost of a pre-payment certificate. All those men with higher medication costs in the UDS arm had purchased pre-payment certificates.

Cost–consequence results

In the primary clinical analysis, UDS is shown to be non-inferior to routine care based on IPSSs over 18 months. From an NHS secondary care perspective, the mean adjusted cost difference showed that UDS cost an extra £216 to achieve a 0.33-point reduction in the IPSS over an 18-month time horizon (*Table 30*). There is no evidence to suggest what society would be willing to pay for a reduction in IPSS.

A key secondary outcome was rates of surgery. No difference was observed in the rates of surgery reported in routine hospital data (30.3% for UDS and 31.3% for routine care). The rates of surgery reported in the CRFs were also similar: 38% and 36% for UDS and routine care, respectively ($p = 0.741$).

TABLE 27 Resource use and costs for the secondary care perspective

Resource category	UDS			Routine care		
	n	Resource use, mean (SD)	Cost, mean (SD) (£)	*n*	Resource use, mean (SD)	Cost, mean (SD) (£)
Inpatient (number of admissions)						
UDS studies[a]	416	0.007 (0.085)	40.83 (481.58)	386	0.005 (0.072)	30.04 (418.83)
TURP	416	0.156 (0.377)	431.31 (1042.91)	386	0.174 (0.393)	475.34 (1081.31)
HoLEP	416	0.031 (0.174)	90.41 (506.04)	386	0.013 (0.113)	36.48 (319.14)
Bipolar resection of prostate	416	0.082 (0.274)	239.20 (834.49)	386	0.078 (0.268)	209.86 (724.72)
BNI	416	0.010 (0.098)	22.00 (224.21)	386	0.005 (0.072)	12.61 (175.90)
Open prostatectomy	416	0.005 (0.069)	4.19 (60.41)	386	0.000 (0.000)	0.00 (0.00)
Urethral stricture surgery	416	0.000 (0.000)	0.00 (0.00)	386	0.003 (0.051)	7.53 (148.00)
Catheter-related procedures	416	0.017 (0.129)	35.51 (306.18)	386	0.010 (0.101)	19.31 (225.12)
Admission with other procedures	416	0.012 (0.109)	23.46 (229.41)	386	0.003 (0.051)	7.30 (143.39)
Prostate, bladder or ureteric problems without intervention	416	0.007 (0.085)	27.54 (460.74)	386	0.013 (0.113)	15.44 (136.28)
Admission without intervention	416	0.010 (0.098)	19.95 (211.91)	386	0.003 (0.051)	7.32 (143.90)
Total inpatient resource use	416	0.34 (0.57)	934.40 (1629.88)	386	0.23 (0.64)	821.24 (1408.73)
Day case (number of attendances)						
UDS studies[a]	416	0.123 (0.370)	37.95 (114.43)	386	0.052 (0.233)	16.02 (72.18)
TURP	416	0.005 (0.069)	7.37 (107.68)	386	0.005 (0.072)	9.63 (133.54)
HoLEP	416	0.007 (0.085)	13.40 (157.37)	386	0.010 (0.101)	18.06 (177.91)
Bipolar resection of prostate	416	0.000 (0.000)	0.00 (0.00)	386	0.010 (0.101)	18.50 (181.16)
UroLift	416	0.005 (0.069)	5.78 (83.32)	386	0.008 (0.088)	9.86 (111.70)
Resection of prostate using VaporTrode (VaporTrode VE-B; Circon ACMI, Stamford, CT, USA)	416	0.000 (0.000)	0.00 (0.00)	386	0.003 (0.051)	4.81 (94.55)
Prostatic artery embolisation	416	0.000 (0.000)	0.00 (0.00)	386	0.003 (0.051)	4.40 (86.45)
Urethral stricture surgery	416	0.010 (0.098)	11.63 (119.62)	386	0.008 (0.088)	8.05 (91.08)

continued

TABLE 27 Resource use and costs for the secondary care perspective (*continued*)

	UDS			Routine care		
Resource category	*n*	Resource use, mean (SD)	Cost, mean (SD) (£)	*n*	Resource use, mean (SD)	Cost, mean (SD) (£)
Catheter-related procedures	416	0.014 (0.119)	4.77 (39.45)	386	0.052 (0.255)	17.23 (84.78)
Other day-case procedures	416	0.019 (0.169)	18.98 (181.81)	386	0.013 (0.113)	11.64 (109.66)
Prostate, bladder or ureteric problems without intervention	416	0.014 (0.138)	4.34 (42.48)	386	0.036 (0.304)	11.91 (96.77)
Day-case attendance without intervention	416	0.024 (0.168)	8.52 (59.81)	386	0.028 (0.182)	10.06 (63.94)
Dialysis sessions	416	0.214 (4.364)	33.43 (681.90)	386	0.000 (0.000)	0.00 (0.00)
Total day-case resource use	416	0.44 (4.53)	148.89 (874.95)	386	0.23 (0.64)	140.18 (428.02)
Outpatient (number of attendances)						
UDS studies[a]	416	0.490 (0.680)	61.51 (85.19)	386	0.122 (0.392)	15.22 (49.06)
Uroflowmetry	416	0.154 (0.439)	19.26 (54.94)	386	0.212 (0.526)	26.42 (65.40)
Micturating cystography	416	0.019 (0.183)	2.41 (22.80)	386	0.021 (0.190)	2.59 (23.76)
Biopsy	416	0.012 (0.129)	3.23 (34.55)	386	0.013 (0.134)	3.48 (36.02)
Cystoscopy	416	0.147 (0.422)	21.23 (61.07)	386	0.114 (0.364)	16.45 (52.58)
Catheter-related procedures	416	0.053 (0.235)	6.63 (29.37)	386	0.054 (0.249)	6.84 (31.34)
Insertion of neurostimulator	416	0.000 (0.000)	0.00 (0.00)	386	0.028 (0.560)	5.62 (110.49)
Nerve conduction studies	416	0.010 (0.120)	1.67 (20.44)	386	0.008 (0.088)	1.20 (13.62)
Ultrasonography (non-radiology)	416	0.005 (0.069)	0.37 (5.65)	386	0.005 (0.072)	0.27 (3.72)
Other procedure not listed	416	0.351 (0.911)	40.85 (115.03)	386	0.329 (1.106)	39.08 (159.47)
No procedure	416	2.063 (2.086)	222.28 (238.13)	386	1.915 (1.857)	200.67 (198.82)
Chemotherapy	416	0.000 (0.000)	0.00 (0.00)	386	0.034 (0.662)	6.02 (118.30)
Radiotherapy	416	0.077 (1.569)	9.88 (201.49)	386	0.000 (0.000)	0.00 (0.00)
Consultation for erectile dysfunction	416	0.000 (0.000)	0.00 (0.00)	386	0.005 (0.072)	0.59 (8.17)
Preoperative clinic	416	0.024 (0.153)	3.73 (23.79)	386	0.028 (0.182)	4.44 (28.31)
Total outpatient resource use	416	3.403.37	393.04 (417.83)	386	2.89 (2.78)	328.90 (357.43)

Resource category	UDS			Routine care		
	n	Resource use, mean (SD)	Cost, mean (SD) (£)	*n*	Resource use, mean (SD)	Cost, mean (SD) (£)
Diagnostic imaging						
Plain radiography	416	0.026 (0.401)	0.79 (9.21)	386	0.008 (0.088)	0.30 (4.15)
Computed tomography	416	0.147 (0.506)	13.66 (47.78)	386	0.135 (0.470)	12.28 (42.30)
Magnetic resonance imaging	416	0.111 (0.395)	16.15 (58.28)	386	0.109 (0.437)	16.87 (66.64)
Ultrasonography (radiology)	416	0.144 (0.575)	7.44 (29.61)	386	0.168 (0.695)	8.75 (35.86)
Positron emission tomography	416	0.002 (0.049)	0.68 (13.78)	386	0.005 (0.072)	1.14 (15.81)
Fluoroscopy/fluoroscopic-guided procedures	416	0.012 (0.109)	1.48 (13.48)	386	0.010 (0.144)	1.29 (17.95)
Echocardiography	416	0.012 (0.109)	0.88 (7.99)	386	0.026 (0.202)	1.84 (14.50)
Total diagnostic imaging resource use	416	0.45 (1.15)	41.07 (100.46)	386	0.46 (1.13)	42.47 (103.61)
Total accident and emergency resource use	416	0.34 (0.84)	63.07 (154.10)	386	0.42 (1.14)	72.14 (189.99)
Total secondary care cost unadjusted	416		1581.23 (2163.23)	386		1409.74 (1692.30)

a   UDS studies here refers specifically to cystometry and pressure–flow studies.

Notes

Resource use and costs are presented disaggregated by procedure for inpatient admissions, day-case and outpatient attendances, diagnostic imaging attendances and accident and emergency department attendances. All costs presented are mean costs discounted at 3.5%, accounting for the 18-month time horizon.

TABLE 28 Resource use and costs for additional secondary care and primary care

Resource category	UDS			Routine care		
	n	Resource use, mean (SD)	Cost, mean (SD) (£)	*n*	Resource use, mean (SD)	Cost, mean (SD) (£)
Additional secondary care						
Inpatient stays at other NHS hospitals (number of spells)	218	0.02 (0.18)	49.05 (377.19)	184	0.02 (0.13)	46.26 (360.33)
Inpatient stays at private providers (number of spells)	217	0.02 (0.17)	46.14 (398.34)	186	0.03 (0.21)	92.56 (589.44)
Outpatient visits at other NHS hospitals (number of attendances)	99	0.09 (0.38)	24.90 (187.63)	92	0.04 (0.21)	8.28 (50.55)
Outpatient visits at private providers (number of attendances)	102	0.03 (0.22)	26.29 (245.75)	97	0.01 (0.10)	1.06 (10.46)
Total additional secondary care resource use	90		119.46 (751.74)	84		112.81 (697.24)
Primary and community health care						
Face-to-face GP appointments (number of contacts)	245	0.69 (1.50)	19.76 (43.28)	218	0.72 (1.79)	20.75 (51.83)
GP home visits (number of contacts)	262	0.00 (0.06)	0.34 (5.53)	233	0.01 (0.13)	0.75 (11.52)
Telephone calls with GP (number of telephone calls)	257	0.22 (0.75)	3.22 (10.91)	224	0.21 (1.10)	3.04 (15.92)
Practice nurse appointments (number of contacts)	256	0.11 (0.47)	0.60 (2.61)	223	0.17 (0.78)	0.95 (4.23)
District nurse visit (number of visits)	262	0.14 (1.20)	5.44 (46.25)	229	0.04 (0.43)	1.69 (16.51)
Community continence nurse (number of contacts)	260	0.04 (0.45)	3.46 (36.98)	233	0.01 (0.15)	1.04 (11.81)
NHS 111 telephone calls (number of telephone calls)	260	0.08 (0.53)	0.94 (6.48)	229	0.02 (0.13)	0.21 (1.60)
Community urology service	258	0.00 (0.06)	0.40 (6.41)	230	0.00 (0.00)	0.00 (0.00)
NHS medication costs	229	13.80[a] (11.02)	101.14 (89.81)	197	13.39[a] (11.48)	96.66 (89.33)
Total primary care resource use	202		127.29 (125.31)	180		120.82 (119.63)

a Mean number of medications.
Notes
Additional secondary care resource use and costs are presented disaggregated by provider (NHS or private) for inpatient admissions and outpatient attendances. Primary care resource use and costs are disaggregated by health-care provider. Discrepancies in costs for what appear to be similar inpatient and outpatient resource use for private providers is due to high cost procedures, which count as one attendance but incur a larger cost. All costs presented are mean costs discounted at 3.5%, accounting for the 18-month time horizon.

TABLE 29 Resource use and costs for the patient perspective (medication, travel and lost earnings)

Resource category	UDS			Routine care		
	n	Resource use, mean (SD)	Cost, mean (SD) (£)	*n*	Resource use, mean (SD)	Cost, mean (SD) (£)
Travel to primary care: car mileage	218	1.64 (4.46)	0.74 (2.01)	192	1.68 (5.16)	0.76 (2.32)
Travel to secondary care: car mileage	156	19.12 (19.02)	21.42 (17.12)	118	17.90 (21.49)	20.67 (19.34)
Personal co-payments for medications	256	13.80[a] (11.02)	5.84[b] (26.76)	232	13.39[a] (11.48)	6.61[b] (29.91)
Over-the-counter medications	262		2.30 (17.37)	230		2.35 (14.09)
Unpaid days off work[c] (*n*)	257	0.68 (3.56)	39.22 (205.29)	228	0.56 (5.71)	32.08 (329.045)
Total personal cost unadjusted[d]	102		104.45 (309.29)	71		45.45 (106.02)

a Mean number of medications.
b Mean cost for personal co-payments.
c Based on an 8-hour working day with paid breaks.
d The total unadjusted personal cost is more than the sum of the unadjusted mean costs quoted in the table. It is calculated from a smaller number of complete cases with a large SD. It is similar to the adjusted costs for perspective 3 in the cost-effectiveness analysis. All costs presented are mean costs discounted at 3.5%, accounting for the 18-month time horizon.

TABLE 30 Cost–consequence results

Variable	*n* (UDS : routine care)	Adjusted UDS, mean (95% CI)	Adjusted routine care, mean (95% CI)	Adjusted difference in means (95% CI)
QALY[a]	416 : 386	1.144 (1.12 to 1.16)	1.138 (1.12 to 1.16)	0.006 (–0.023 to 0.035)
IPSS score[b]	340 : 329	12.61 (7.92)	13.11 (7.86)	–0.33 (–1.47 to 0.80)
NHS secondary care costs[c]	416 : 386	1602.51 (1426.45 to 1778.57)	1386.80 (1203.92 to 1569.67)	215.72 (–39.95 to 471.38)
NHS costs[c]	416 : 386	1852.51 (1655.90 to 2049.12)	1710.64 (1503.25 to 1918.03)	141.87 (–145.56 to 429.29)
Patient costs[c]	416 : 386	103.40 (68.09 to 138.70)	59.92 (26.13 to 93.72)	43.47 (–5.77 to 92.71)

a Adjusted for centre, baseline IPSS score and baseline utility.
b Unadjusted means (SDs) shown. Difference in means adjusted for centre, baseline IPSS score and baseline utility.
c Adjusted for centre, baseline IPSS score and baseline resource use.
Notes
Cost–consequence analysis of UDS vs. routine care over an 18-month time horizon. In this case, the time horizon is equal to the duration of the follow-up period. The maximum QALYs available over an 18-month period is 1.5.

Men in the UDS arm had a similar number of adjusted mean QALYs (1.14, 95% CI 1.12 to 1.16), compared with the routine care arm (1.138, 95% CI 1.12 to 1.16) (see *Table 30*). The difference of –0.006 (95% CI –0.023 to 0.035) is equivalent to 0.45 fewer days of the best imaginable health over the 18-month time horizon for those men in the UDS arm.

From the wider NHS perspective, the UDS care pathway cost an additional £142 (95%CI –£146 to £429) compared with routine care. In the analysis from the patient perspective, UDS was estimated to be £43 more expensive than routine care (95% CI –£6 to £93).

Cost-effectiveness results

Cost-effectiveness analyses show that the UDS pathway is more expensive, with similar QALYs over an 18-month time horizon compared with the routine care pathway for all three perspectives (secondary care, wider NHS and patient's perspective). From the primary perspective of secondary care, we estimated the ICER for UDS over routine care to be £35,554 per QALY over 18 months, with an INMB of –£94 for UDS over routine care at a societal willingness to pay of £20,000 per QALY (95% CI –£719 to £530). For the wider NHS perspective, we estimated the ICER for UDS over routine care to be £23,378 per QALY over 18 months, with a respective INMB of –£21 (95% CI –£658 to £617) at a willingness to pay of £20,000 per QALY. The estimated ICER for UDS compared with routine care for the patient's perspective was £7171 per QALY, with a respective INMB of £78 (95% CI –£505 to £661) at a willingness to pay of £20,000 per QALY. There is significant uncertainty (wide CIs) surrounding the cost, QALY and INMB estimates.

Figure 12 illustrates the CEACs for the three different perspectives. The probability that the UDS pathway is the cost-effective management compared with routine care is 0.38 at willingness-to-pay threshold of £20,000 and 0.47 at a willingness-to-pay threshold of £30,000. The probability of UDS being cost-effective compared with routine care rises to 0.6 when the societal willingness to pay (i.e. cost-effectiveness threshold) is £100,000 per QALY.

Sensitivity analyses

All sensitivity analyses (*Table 31*) supported the findings of the main analysis, which showed that UDS is both costlier than and similarly effective as routine care.
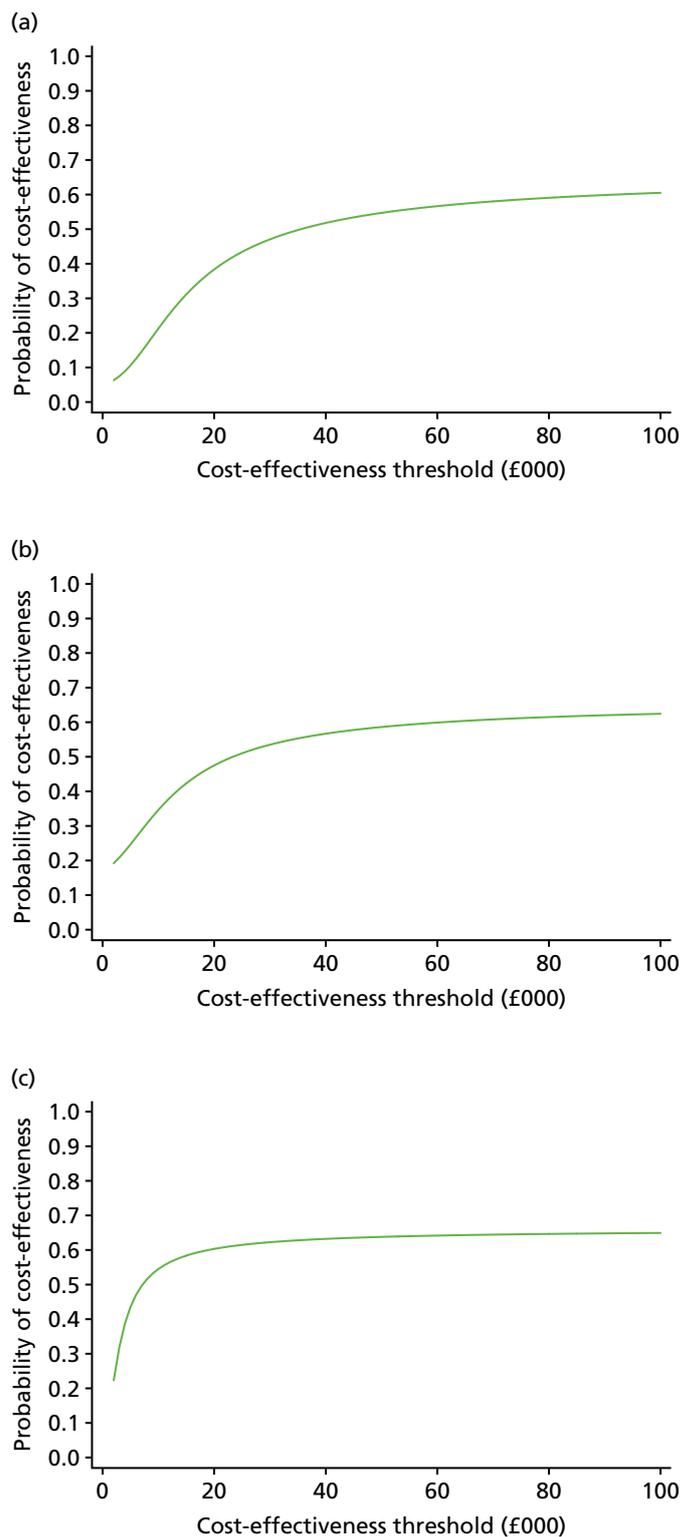
FIGURE 12 Cost-effectiveness acceptability curves for (a) secondary care; (b) the wider NHS perspective; and (c) the patient's perspective. CEACs illustrate the probability that the UDS pathway is the cost-effective option compared with routine care.

TABLE 31 Sensitivity analyses

Arm	n	Adjusted costs (£), mean	Adjusted QALYs, mean	Incremental costs (£), mean (95% CI)	Incremental QALYs, mean (95% CI)	ICER (£/QALY)	INMB (£) at £20,000/QALY
Upper estimate of reference cost used							
UDS	416	1847.02	1.144			40,627.86	
Routine care	386	1600.52	1.138	246.47 (−46.13 to 539.13)	0.006 (−0.023 to 0.035)		−125.16 (−765.21 to 514.90)
Lower estimate of reference cost used							
UDS	416	1231.24	1.144			24,886.57	
Routine care	386	1080.26	1.138	150.97 (−51.77 to 353.72)	0.006 (−0.023 to 0.035)		−29.64 (−634.37 to 575.09)
0% discount rate used for both costs and effects							
UDS	416	1612.03	1.157			37,439.11	
Routine care	386	1394.73	1.151	217.31 (−39.83 to 474.45)	0.006 (−0.023 to 0.035)		−101.22 (−733.85 to 531.41)
5% discount rate used for both costs and effects							
UDS	416	1597.66	1.138			34,818.71	
Routine care	386	1382.67	1.132	214.99 (−39.99 to 469.98)	0.006 (−0.022 to 0.035)		−91.50 (−712.28 to 529.28)
Excluding participants who did not have a complete 18 months of follow-up							
UDS	384	1642.06	1.142			30,409.32	
Routine care	357	1363.47	1.133	278.59 (8.67 to 548.52)	0.009 (−0.021 to 0.039)		−95.36 (−743.19 to 552.46)
Conducting the analysis censoring everyone at 15 months							
UDS	416	1496.10	0.96			25,928.45	
Routine care	386	1265.05	0.95	231.06 (−18.86 to 480.98)	0.009 (−0.014 to 0.032)		−52.83 (−568.56 to 462.90)
Wider NHS perspective without private care resource use							
UDS	416	1811.45	1.143			32,959.71	
Routine care	386	1611.43	1.138	200.03 (−66.06 to 466.07)	0.006 (−0.023 to 0.035)		−78.65 (−708.78 to 551.49)
Costs of UDS tests identified in CRF but not routine care							
UDS	416	1644.18	1.143			41,722.29	
Routine care	386	1391.03	1.138	253.15 (−2.41 to 508.71)	0.006 (−0.023 to 0.035)		−131.80 (−756.00 to 493.00)

Notes
All sensitivity analyses presented are perspective 1, the secondary care perspective, except sensitivity analysis 8, which presents the wider NHS perspective without private providers. The time horizon for all estimates is 18 months, except for sensitivity analysis 4, which has a time horizon of 15 months.

Discussion

Overview

We performed a within-trial, regression-based, cost-effectiveness analysis of an invasive test within a diagnosis pathway for men presenting with LUTS. The results of the economic analysis suggest that a pathway with UDS is more costly than routine care and produces similar QALYs over an 18-month time horizon.

Strengths

Our results are robust to sensitivity analyses, testing the effect of using upper and lower quartile reference cost valuations, changes in the discount rate from 0% to 5%, censoring all participants at 15 months, or excluding those participants without complete routine data follow-up.

Obtaining information about secondary care resource use directly from hospital informatics departments had several advantages. First, this method allows us to assume that the data received are complete, because they were extracted from hospital routine reporting systems. What is more, any missing data are most likely to be due to administrative error and could therefore be considered missing at random (i.e. missing for a reason unrelated to patient outcomes). Second, we were able to obtain detailed data at the level of FCEs. The data included the treating specialty, multiple diagnosis and procedure codes, and information about additional services used, such as diagnostic imaging or chemotherapy treatment. The richness of this data set allowed us to confidently select resource use related to LUTS to include in our evaluation. Third, the data in question are routinely collected by hospitals. Extracting the data in this way meant that the task was not unnecessarily duplicated by research nurses. Finally, the information was obtained from independent information analysts who were unaware of participants' assigned treatment and were not involved in the design or implementation of the trial; therefore, the reported data should be free of any bias.

Limitations

There are limitations to using routine hospital data in clinical trials. In this trial, we approached only the main treating hospitals for their routine data; it is therefore very likely that operations or investigations carried out at other hospitals may be missing. At many sites, outpatient clinical coding was performed by clinic staff and not professional clinical coders. It is possible that some of the UDS tests may not have been coded correctly by clinic staff and, therefore, may have been classified simply as a urology outpatient appointment, without a procedure code identifying the UDS procedure.

We found that using routine data led to an underestimation of the number of UDS tests carried out in the UDS group. The number of operations was also underestimated from the routine data compared with the CRF data; however, additional operations at other NHS or private providers were identified and costed in the wider NHS perspective. We tested the impact of missing UDS test records in a sensitivity analysis. Adding this cost increased the adjusted mean cost in the UDS arm by £37 and increased the ICER to £41,722 per QALY gain. The ICER is particularly sensitive to changes in cost because of the small difference in QALYs.

It should be noted that underestimates for UDS testing and surgery do not invalidate using routine data as a methodology. UDS testing and surgical outcomes were key trial variables and it is probable that research nurses paid particular attention to collecting these data. Further work is necessary to investigate the reliability of using routinely collected data as we move towards efficient trial designs.

Two further sensitivity analyses specifically addressed the limitations of using routine data in a time-constrained situation. We addressed administrative censoring by excluding all those participants with incomplete follow-up (32 men from the UDS arm and 29 men from the routine care arm). UDS remained costlier and similarly effective, with a similar ICER to the main analysis (£30,409/QALY and £35,554/QALY for the UDS and routine care arms, respectively). Censoring all participants at 15 months effectively shortened the time horizon to 15 months. Again, UDS was costlier, with a cost difference of £231.06 for the sensitivity analysis compared with £215.72 for the main analysis, with a similar QALY difference of 0.009 (0.006 for the main analysis) and an ICER of £25,928 per QALY. These sensitivity analyses support the usefulness of routine data in NHS clinical trials.

A further limitation of our analysis is the low levels of resource use data completion from participants, particularly in relation to the patient's perspective evaluation. To include as many patient-reported data as possible, we assumed that any questions unanswered in a returned questionnaire reflected zero resource

use for those items. Despite this assumption, there were insufficient complete data from the resource use questionnaire to accurately cost patient costs, specifically for travel (other than car travel) and the amount spent on incontinence pads. We had originally intended to explore this assumption regarding unanswered questions in a sensitivity analysis; however, the lack of complete data across the resource use questionnaire made this sensitivity analysis impossible. More weight should therefore be given to the primary analysis with the secondary care perspective.

The resource use questionnaires were administered every 6 months. Men were asked to recall their resource use over the 6-month duration, which is a potential limitation. It is difficult for most people to recall all their health resource use over a 6-month period. We supplied men with a resource use diary to be used as an aide memoire. Previous research has shown that this can improve patient recall and accuracy of resource use reporting;[46] however, in this trial it did not seem to help completion rates.

We costed medications based on the number of medications and an average 'actual cost' per GP prescription, which is derived from net ingredient costs minus the assumed average discount negotiated by the NHS. Although information was obtained from men regarding actual medication use, we made a decision not to use this method, given that it is very labour intensive to cost, and prone to missing data and research has shown that, when the main intervention is not pharmacological, the method of costing medications has minimal effect on the overall estimated cost-effectiveness of the intervention.[47]

The EQ-5D-5L instrument, with the exception of the baseline questionnaire, was included in the resource use questionnaire, which was separate to the symptom-based questionnaire that contained the IPSS primary outcome. This possibly contributed to the low numbers for whom a QALY could be created for. It may therefore be advisable that in future studies the EQ-5D-5L is included in the same questionnaire as the primary outcome measures.

Future work
Urodynamics is not the sole test used in determining whether or not a man with LUTS should be offered surgery. There are a number of possible treatments that men may be offered based on these collective tests: watchful waiting, pharmacological management, conservative treatment, minimal invasive techniques (e.g. UroLift), prostate artery embolisation or surgery (e.g. TURP) and greenlight laser. The most cost-effective diagnostic strategy within this type of diagnostic pathway could be assessed using a decision model.

Conclusions

The results of the economic analysis suggest that including UDS in the assessment of men who present with LUTS, compared with routine care, leads to higher costs and similar QALYs from all three perspectives (i.e. NHS secondary care, a wider NHS perspective and patient perspective). There is great uncertainty surrounding the estimates presented, as evidenced by the very wide and overlapping CIs. However, all of the sensitivity analyses support the main results and demonstrate that they are robust to the limitations of using routine data. CEACs demonstrate that the probability that UDS is cost-effective compared with routine care is 0.38 at a societal willingness to pay of £20,000 per QALY, rising to a probability of 0.6 if the willingness to pay is increased to £100,000 per QALY.

Chapter 5 Qualitative evaluation

Some of the work reported in this chapter has been reproduced from Selman *et al.*[48] © 2017 Selman LE, Ochieng CA, Lewis AL, Drake MJ, Horwood J. *Neurourology and Urodynamics.* Published by Wiley Periodicals, Inc. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. See: https://creativecommons.org/licenses/by-nc/4.0/.

Introduction

The value of integrating qualitative methods in randomised controlled trials includes exploring the views and experiences of trial participants and those recruiting to the trial.[49,50] In UPSTREAM, qualitative methods were used to illuminate the perceived value and acceptability of invasive UDS testing for male BOO and its impact on clinical decision-making, and to explore any barriers to its uptake outside the trial. Patient and clinician views of UDS (including its acceptability and tolerance among patients) are critical to inform the implementation of trial findings and the use of UDS in practice. Some of the results presented here have already been published.[48]

The specific aim and objectives of the qualitative component of the trial were as follows.

Aim
To understand patients' and clinicians' views and experiences of UDS testing for male BOO and BOO surgery.

Objectives

- To understand patients' and clinicians' experiences of the trial, including their experiences and views of UDS/routine care (e.g. the acceptability and feasibility of UDS).
- To investigate patients' and clinicians' decision-making regarding surgery for male BOO.
- To understand barriers to and facilitators of UDS.
- To explore the information and support needs of patients and clinicians in relation to UDS and BOO surgery.
- To investigate patients' and clinicians' experiences, attitudes and opinions regarding male BOO surgery and recovery.

UPSTREAM participants' views and experiences

Methods

Trial design
In-depth semistructured interviews were conducted with patients enrolled in UPSTREAM,[19] at 26 secondary care urology centres across England. Interviews allow for the exploration of complex and sensitive issues, allowing participants to engage in a dialogue in their own language and drawing on their life experiences to cover the issues that are important to them.

Recruitment and sampling
All trial participants had consented at baseline to being approached by a qualitative researcher. Purposive sampling was used to select a maximum variation sample in relation to (1) trial arm; (2) trial centre; (3) sociodemographic and clinical variables, age, ethnicity and socioeconomic status; and (4) baseline total

IPSS.[51] Socioeconomic status was estimated using the Index of Multiple Deprivation decile,[52] which ranks areas in England from most deprived (score of 1) to least deprived (score of 10); we sampled across three categories (high, 1–4; medium, 5–7; low, 8–10). The IPSS assesses seven symptoms: incomplete emptying, frequency, intermittency, urgency, weak stream, straining and nocturia; each is scored 0 (best) to 5 (worst). Total IPSS at baseline was categorised as high or low using the American Urological Association Symptom Index classification: ≥ 20 indicates high symptom burden (severe symptoms) and ≤ 19 indicates low symptom burden (mild or moderate symptoms).[51] For the post-surgery patients, their most recent post-surgery IPSS was also considered during sampling when available.

Patients were recruited for interview at two time points: either 1–8 weeks after the consultation in which their treatment had been decided and recorded, or 6 weeks to 4 months after receiving surgery for LUTS. This was to capture variation along the treatment pathway and in treatment types (surgery vs. non-invasive treatment).

Interviews

Participants in the local Bristol area were interviewed either face to face or by telephone, depending on their preference. Participants in other geographical areas were interviewed by telephone. All interviews were conducted by Cynthia A Ochieng or Lucy E Selman, both experienced and trained qualitative researchers with a background in health services research. Neither researcher works clinically in urology, nor was previously known to participants.

Interviews followed topic guides developed at the start of the trial based on the literature, the qualitative trial aims and input from the TMG, including patient representatives (see *Appendix 5*). The topic guides were devised to ensure that the primary issues were covered across all interviews, but did not dictate data collection. The topic guide incorporated considerable flexibility to enable participants to introduce new issues unanticipated by the researchers. Each topic guide was piloted with two men with LUTS (data not included in analysis), and refined prior to use. The topic guides were further refined as data collection progressed, in response to emerging findings.

In telephone interviews, consent was taken verbally prior to the interview. In face-to-face interviews, participants gave written consent. Field notes were taken during and after the interviews and integrated into the analysis. All interviews were digitally audio-recorded, professionally transcribed verbatim, anonymised and checked for accuracy prior to analysis. Transcripts and field notes were imported into NVivo V10 software (QSR International, Warrington, UK) for analysis.

Analysis of the interviews

Analysis was conducted in parallel with data collection, with findings from early analysis informing later data collection in an iterative process, and sampling continued until no new themes emerged from the analysis by the end of data collection.[53] Thematic analysis[54] was used to analyse the data, following six steps:

1. Familiarisation with the data through detailed reading.
2. Two different members of the qualitative research team independently constructed draft coding frames using line-by-line coding of a sample of the data. A combination of deductive coding, based on the aims of the trial and the topic guide, and inductive coding, identifying themes in the data, was used.
3. The qualitative research team met to compare the coding frames and decide on the final draft of coding frame to be used.
4. Themes were defined (in NVivo) and the coding frame applied to the data by one researcher.
5. The researcher met with the rest of the qualitative team as coding progressed, to review and discuss emerging findings.
6. A narrative was drafted based on the analysis, paying attention to patterns and deviant cases, with input from the qualitative team (LES, CAO, JH, CC). Emerging findings were also discussed at TMG meetings to refine the topic guide and qualitative analysis and to inform trial conduct.

Separate coding frames in the same NVivo project were created for pre-treatment and post-surgery patient data sets. Clinician interviews were analysed in a separate NVivo project. Analysis initially focused on each data set separately, and then considered and integrated findings from the patient and clinician qualitative data sets.

Results

Trial participants

In total, 50 interviews were conducted, comprising nine patients who declined to participate in the trial and 41 trial participants. The trial participants interviewed were from diverse locations in England (25 pre treatment, 16 post surgery) and ranged in age from 52 to 89 years (*Table 32*). Fifteen participants had a high symptom burden at baseline. Twenty-five participants had experienced UDS testing.

TABLE 32 Characteristics of interviewed patients

Participant characteristics	Pre-treatment patients (*n*)	Post-surgery patients (*n*)
Age group (years)		
51–55	1	1
56–60	1	3
61–65	7	2
66–70	7	3
71–75	6	2
76–80	2	3
81–85	1	0
86–90	0	2
Trial arm		
UDS	13	9[a]
Routine care	12[b]	7[c]
Treatment decision		
Conservative	13	N/A
Surgery	12	
Time since surgery (days)		
Median	N/A	91.5
Range		48–463
Centre[d]		
1	4	1
3	0	3
4	1	0
5	1	0
7	0	1
8	1	2
10	1	1

TABLE 32 Characteristics of interviewed patients (*continued*)

Participant characteristics	Pre-treatment patients (*n*)	Post-surgery patients (*n*)
11	2	0
13	0	1
14	0	1
15	1	2
16	3	0
17	3	0
18	1	0
19	0	1
20	4	1
21	0	1
22	1	0
23	1	0
24	0	1
25	1	0
Deprivation decile[52]		
High (1–4)	11	4
Medium (5–7)	9	9
Low (8–10)	5	3
IPSS symptom burden (baseline)[51]		
High ($\geq$ 20)	8	7
Low ($\leq$ 19)	17	9
Ethnicity		
White British	23	12
Asian/British Asian	1	1
White American	0	1
Iranian	0	1
Afro-Caribbean	0	1
Not given	1	0

N/A, non-applicable.
a One of these men was randomised to the UDS arm but did not receive UDS as he was on holiday and then had surgery scheduled.
b Four of these men reported that they had received UDS prior to joining UPSTREAM, as part of routine care earlier in their disease journey.
c One of these men had UDS as part of routine care, after specifically requesting it.
d Centre names have been replaced with numeric identifiers for the purpose of reporting.

Among pre-treatment trial participants, seven agreed to surgery after receiving UDS during UPSTREAM, five agreed to surgery after non-invasive testing (i.e. routine care), six agreed to conservative treatment after UDS during UPSTREAM (although one man was undecided) and seven agreed to conservative treatment after non-invasive testing.

Verbatim quotes are used below to illustrate emerging themes from the analysis, with PT denoting pre-treatment trial participants and PS post-surgery trial participants.

Views and acceptability of UPSTREAM among decliners

Seven interviews were conducted with patients who declined to take part in the trial. Reasons given for not taking part in UPSTREAM included patients wanting to be allocated to UDS (or not), because of their own opinion or that of their consultant, and not wanting to leave the choice of assessment to randomisation. Reasons for not wanting to be randomised to receive UDS included not wanting UDS unnecessarily, fear that having UDS might delay or lead to surgery and not liking the idea of having an invasive UDS test. One decliner had a bad previous experience with UDS and did not want to have it again. Others reported not having the time owing to work constraints or other commitments.

> *I thought, well [–] I actually – I would rather either have or not have a UDS test based on what the consultant, in discussion with me, thinks is necessary or advisable. I didn't just to [–] want to randomly either have it or not have it because I'd just randomly been selected for one group or another.*
>
> *Decliner 1*

> *I've already – I feel that this is, um, a step too far. If I can go in and have the operation without all this nonsense, um, and I'm sorry, I used that word [laughs] but, um, I – I just felt that the sooner I get it over and done with the better.*
>
> *Decliner 4*

> *I felt that, although he [consultant] explained it [surgery] well, he explained it very lightly but, oh yes, I had to go in overnight, er, and, alright, I – I had a pipe up me, my pipe, pipe, er, and I would be alright. But when I read the [–] er [–] booklet that he gave me, I realised that it was a lot more complicated than that, and that I would be out of action for anything up to a month . . . And I could be bleeding, I could have catheters again, so I decided at my age of 85 that it wasn't worth doing. I'll get up, er, more frequently at night, thank you.*
>
> *Decliner 3*

Pre-treatment and post-surgery trial participants

Participants' initial impressions of UPSTREAM and reasons for taking part

All trial participants interviewed were happy to be approached and take part in UPSTREAM. Participants' main motivations to take part were the altruistic belief that taking part would benefit others or the view that they would benefit directly from participating. Some participants were favourably disposed towards research that aimed to improve practice and wanted to help for this reason. Others believed that they would receive quicker and better care and treatment if they were part of the trial, or wanted to know more about their condition and felt that taking part would help with treatment decision-making. One participant wanted to avoid surgery and felt that the trial may provide alternative options:

> *Like treatment for cancers and all sorts of things, someone has to try the drugs at some stage and whatever else you know surgery, it's all in the early days you know done without anaesthetic and that, I wouldn't have wanted that, but um if nobody does it, we'll not move forward.*
>
> *PS8, routine care*

> *I want to help other guys in this situation, and if the information I pass over together with other people's helps to improve um, quality of life for people, then yeah I'm more than happy to get involved.*
>
> *PS9, UDS*

I thought maybe it would speed things up and people would be more aware of it, you know. And not only that, you know, if someone else is watching what's going through the process, through the back door, then maybe I'll get a better service . . . I think if I wasn't on the UPSTREAM programme I'd be just, you know, a Joe Bloggs and a number.

PT17, UDS

Participant understanding of randomisation

Trial participants reported different levels of understanding of the trial aims, the randomisation of participants into two arms, why this was done and what was involved in each arm. Most participants had a clear understanding of what randomisation meant and what the two arms entailed. However, some understood only that they would be randomised and not what being in one of the two arms entailed:

Well as far as I understood it was to see, err, whether actually doing the UDS extra part of it actually improved the diagnosis of the problem . . . I think that's what I understood by it.

PS1, routine care

Well, as I said, it was random anyway. You couldn't opt to go in one group or the other. It was a random selection. Erm, for the second test to prove straightaway that you needed surgery, to me is a better idea.

PT4, UDS

I didn't know what to make of it because I don't know if I . . . if I err . . . I didn't know whether . . . what was going to be done to the other group was, was something good, you know. I never . . . I don't think from the start I had a clear understanding of what the difference was, and when I was told that, you know, I would be in the passive group or whatever the proper term would be.

PS4, routine care

The acceptability of randomisation and equipoise/arm preference

Participants were happy to be randomised and accepted that it was part of the trial design. Although some participants from the routine care arm said that they would have been happy to have been allocated to either group, most participants (from both arms) did express a preference for one arm or the other. Reasons for wanting to be allocated to the invasive UDS testing arm included believing that the UDS arm would be more effective, in that it would provide more complete and accurate information to inform treatment decision-making:

Well, I understand that that's the way trials work. You know, it's got to be randomised, and you have to accept whatever arm of the trial you are on . . . And, and I don't have, you know, I had no feelings one way or another. I think, you know, I understand how clinical trials are done, so, so I think, you know, if you wanna take part, you have to accept whatever happens, whatever group you get assigned to is absolutely fine.

PT5, routine care

I was bitterly disappointed actually that I wasn't in the [UDS arm] . . . It would have been more helpful . . . I think it would have been more interesting for me as well . . . just finding out about how your body works to a certain extent. It's more you know. The more tests the better, I say – to a certain extent, because you know exactly what's going on then.

PS5, routine care

[You] would only get more definitive answers from the UDS test . . . I must admit if I hadn't been selected for that, I think I might have been a bit disappointed . . . clearer set of information on how to make the right decision to operating or not . . . much more accurate assessment of what the effect of surgery would be.

PT10, UDS

What improvements could be made to UPSTREAM?

Although most participants were happy participating in the trial, reported good experiences and did not have any recommendations for improvements, a minority of participants did raise issues. A few participants felt that the questionnaires were restrictive, not allowing a true reflection of their experience. Participants welcomed the interviews as an opportunity to explain, in more depth, their experiences:

> *Sometimes they're looking for, are perhaps looking for black and white answers, whereas, er, not really an opportunity to explain what exactly what has happened.*
>
> *PS12, UDS*

> *I've had the opportunity to take part in a conversation like this which is important, not just for you, but for me.*
>
> *PT25, routine care*

Lack of contact at the start of the trial and having to wait to see a consultant or have tests was a problem for some participants. An accurate timeline for what to expect when consenting to the trial would have been appreciated:

> *I probably haven't treated it [the trial] quite with the respect it deserves, 'cause I got a little bit anti when I was rolling about here and seemed to be waiting an awfully long time for somebody to do something about it rather than ask me questions about it . . . the only thing annoys me at times is I'd rather somebody said to me you, you know . . . you could be waiting 12 or 18 months rather than 12 weeks, because you get quite excited about, it's gonna happen, and then it doesn't. And also when you leave the hospital, they tell you, you know 'Oh, you'll see the consultant in 2 weeks' . . . and it turns out to be nearly 3 months. I'd rather be told up front what it's likely to be than, um, wildly inaccurate figures.*
>
> *PS12, routine care*

Acceptability and experiences of urodynamics

All 25 men who had undergone UDS testing (either during UPSTREAM or earlier in their illness journey) reported that it was acceptable, despite any discomfort or other issues they might have experienced. Of the 16 men who had not had UDS previously, 10 said they would have been happy to have it if needed:

> *I'm used to being prodded and god knows what else. I don't take any notice of that . . . The more tests the better I always say . . . you know exactly what's going on then.*
>
> *PS5, routine care*

Of the remaining six men, four said that they would have had it if needed, but would have been apprehensive because of its invasive nature and were glad it had not been required. Two men said that they would want more information about the test and its purpose. One man had declined to have UDS because a prior cystoscopy had found a narrow constriction and he did not want to risk any damage that might delay his planned surgery.

The main reason for wanting UDS was the perceived accuracy of the test and the desire to have maximum information to inform the treatment decision:

> *I like answers . . . and we got the answers. So if I hadn't had the test done, then we're just left with a load of symptoms which could be caused by this or that. Erm, so I mean, in terms of, erm, the invasion of the test, it's . . . not a problem . . . totally acceptable.*
>
> *PT20, UDS*

We did not find any evidence of associations between age or symptom severity and the acceptability of UDS.

Experience of urodynamics: pain, discomfort and infection

In general, UDS was well tolerated by the 25 participants who had experienced it. However, there was variation in how uncomfortable men found the procedure. Seven men reported that it was at least a little painful, with some experiencing severe pain, but they, nevertheless, thought UDS was acceptable:

Having a tube stuck inside me and afterwards was 10 out of 10 pain . . . It was awful. But the people I dealt with were very good . . . Yes [acceptable], no problem. Obviously embarrassing, uncomfortable, not very pleasant. Erm – well, I mean, I had the choice of having the test done or not, so, you know, I read about it, it sounded horrible, but, erm – I am quite a positive person. I want to live the best. I've got a family to look after.

PT6, UDS

Five men reported that the procedure was pain free:

I was quite surprised, actually. Especially at the, at the rear end, I thought, 'Oh, there might be a pain involved here'. But there wasn't. I had no pain at all. I just – it was, it was better than what I thought it was going to be.

PT11, UDS

Urodynamics was sometimes perceived as easier and less uncomfortable than uroflowmetry, because UDS does not involve having to wait with a full bladder:

When they put liquids into you, [it] was more uncomfortable . . . easier . . . It's able to fill your bladder through your penis, so that they can measure the flow . . . That was a bit embarrassing, but like I said, I found it better because I didn't have to retain anything. So to me . . . although embarrassing, [UDS] is a better system [than flowmetry].

PT4, UDS

Eight participants reported short-lived negative after-effects of UDS: stinging when urinating, a small amount of bleeding, a UTI or disrupted flow/urgency. However, despite these issues, the men said they would willingly have the test again if needed.

Experience of urodynamics: embarrassment

A minority of participants (7/25) reported that having UDS was embarrassing, because of either its intimate nature or not being prepared for its effects:

The final pee that I was having was with a catheter up my urethra. Now, how you're supposed to pee into a jug when you've got a catheter up you, I do not know. But it went all over the wall and the floor, which embarrassed me. If they'd said, 'Look, this is what's going to happen, don't worry about it' . . . one of the nurses had to clean it all up when she came back in.

PT10, UDS

One patient reported that he found UDS less embarrassing than a rectal examination:

. . . it's not very nice for them anyway, and certainly not nice for you either.

PT4, UDS

The degree to which UDS assessments (and other tests for LUTS) were perceived as embarrassing, depended in part on the level of privacy available, including the number of people in the room during the test, room location and size (a larger room near a corridor was more socially awkward). Patients preferred as few people present in the examination room as possible during the test, that staff were introduced to them and that they knew their roles:

PT11, UDS: *It was all right. Embarrassing [laughs]! Stood there in front of all the people! . . . There were a couple of students! [Laughs] A couple of nurses! [Laughs].*

Researcher: *Were you asked about whether you minded them being there, or . . . ?*

PT11, UDS: *Yeah, yeah, [the doctor] asked me if I minded if they'd be there . . . I didn't mind. Yeah, a little bit embarrassing.*

Who performed the UDS test could also be a concern and a cause of embarrassment. Two men mentioned the sex of the person performing the test in relation to their embarrassment. One would have preferred a woman to have carried out the test:

PT4, UDS: *Well, it's not very nice for another man to play with your . . . Well, not play with you but actually touch you, they're doing things. Then to stand in front of women as well . . .*

Researcher: *Oh, there were women in the room?*

PT4, UDS: *Yeah, well, I was asked about . . . One was a trainee nurse and one was a senior nurse. So to be quite honest, [–] to me, it's easier for me to be naked in front of a woman than what it is a man.*

Another said that his test had been performed by a female nurse and initially he thought he would have preferred a man. However, by the end of the test he thought it was better to have it performed by a woman. Despite any embarrassment, men still found UDS acceptable.

Participants' perceptions of the value of urodynamics

As previous results suggest, UDS was primarily valued by participants for the additional insight it gave them and their clinicians into their LUTS:

I was glad that it was clear result in itself . . . it was effective in actually showing, illustrating the problem and the extent of the problem . . . it was effective in actually getting something sorted.

PT22, UDS

Urodynamics was often the last in a series of assessments participants had received for their LUTS. Many patients felt satisfaction that having UDS meant that they had received all the investigative tests available and, therefore, had all possible facts regarding their condition:

You're having a test that was different from what erm, you'd experienced before . . . there was a sort of finality about that, you know. You felt that, you know, OK [–] there's nowhere else to go, you know.

PT11, UDS

Three patients reported that they found invasive UDS testing interesting or engaging as they learnt about the cause of their LUTS. UDS was perceived as more informative than other tests, providing a more accurate account of the cause of LUTS:

I done the flow rate thing, the UDS was I think far more accurate . . . I thought it was a very clever test . . . You know, I think it showed a lot more . . . that other test was probably one of my bad days, because the growth of a prostate outside the bladder acts like a ball-valve so it shuts my flow off . . . Yeah, I think it was a good test.

PT15, UDS

For some patients, the results of UDS played an important role in their treatment decision-making [see *Participant perspectives on the role of assessments (including urodynamics) in lower urinary tract symptoms treatment decision-making*]:

. . . although they weren't pleasant, you feel as though at least everybody's had a go, and they've done as much as they can. And I think when [consultant] sort of turned round and sort of said,

'Well, that confirms everything' . . . and he's not made any secrets of what he thought, so you knew what he was talking about.

PT13, UDS

Information provision for urodynamics

Although, overall, men's experiences of care were positive, there were instances along the diagnostic pathway when participants reported that they had been inadequately informed. Four patients felt that they had not been fully informed on the process of UDS in advance, such as what it would involve, the need to urinate and possibly spray, or the risk of developing a UTI afterwards:

They put a catheter in, and I told them, the nurse who was doing it . . . 'When you put that catheter in', I said, 'I know for a fact, by at least tomorrow, I'm going to have an infection'. – 'No, everything's sterilised. You won't have an infection'. True enough, I had an infection, and it was a Saturday the next day. My doctor was closed.

PT11, UDS

Although most participants reported being given leaflets about the test prior to having it, one patient did not know he was having UDS until he arrived at the hospital on the day of the test. Another participant described the importance of patients feeling fully informed while awaiting their tests:

Very clear information is very important. 'You'll be sitting here; you'll be sitting here for an hour. It'll be that length of time before we will call you, or this will take, given the number of people we've got, I expect it'll be this.' If it's wrong, it's wrong, it doesn't matter, but actually giving clear and confident messages allows people to feel that they're not completely out of control . . . they haven't been forgotten in the corner.

PT22, UDS

There was variability in participants' reports of how and when their test results were explained to them and the adequacies of the explanations they received. Men had the results of UDS explained to them during the test by the technician or nurse undertaking it, from a doctor straight after receiving the test, or at a separate appointment with a doctor a short time later. When test results were available and discussed with a clinician immediately after the test, this was appreciated:

I had an instant diagnosis, saying that I needed an operation . . . I weren't expecting that. I was expecting to have a letter a week later saying, 'We think this, that and the other'.

PT4, UDS

[Getting] the information on the day, obviously that's better for everybody, you know. I think it takes the angst out of it really.

PT1, routine care

Most men were satisfied with the explanation of test results, reporting that the explanations received were excellent or adequate:

So, I could understand, you know, what results they were looking for and they actually showed me the graph results. So erm, that was quite interesting. That was explained.

PT13, UDS

Researcher: *Having done the different tests, how were the results explained to you?*

PT6, UDS: *Yes, very well . . . The nurse explained it to me as well, and then I came back to the doctor and he went through it with me.*

Two men reported poor explanations of UDS results. One reported a problem with the results of the UDS being explained in a rushed manner and immediately after the test:

> *The consultant was in a hurry, so he started trying to explain the readings off the test on the computer to me while I was still standing there in a gown with a catheter in. And I said, 'I'm sorry. I can't take this in now. Can you wait until I've got dressed?' . . . The only thing he [doctor] said to me was this thing about, 'There's an awful lot of pressure going in, and not a lot of urine coming out.' Erm, so I saw, you know, the computer monitor and one or two printouts, but I didn't ask. I wasn't shown those.*
>
> *PT10, UDS*

Participant perspectives on the role of assessments (including urodynamics) in lower urinary tract symptoms treatment decision-making

Participants valued assessments for their information and reassurance: assessments answered questions they had, helped clarify the cause of their LUTS, validated what they and their clinicians had already suspected, and provided reassurance. For some participants, assessments provided new information, confirming either that they had a problem that needed to and could be treated, or that they would not benefit from surgery:

> *I didn't realise that I had a flow rate problem until I had this flow test done.*
>
> *PT13, UDS*

> *The results meant that I didn't have a problem that I thought, you know, we originally thought we had, because they wondered if I had a narrowing of the tube through the, through the prostate because of the enlargement. But there was no narrowing, and it was perfectly all right.*
>
> *PT7, routine care*

Urodynamics, in particular, was valued for its perceived accuracy, for example in showing whether or not the bladder outlet was obstructed and therefore whether or not the patient would benefit from BOO surgery:

> *It [UDS test] decided me . . . it determined if I'm having the operation or not. They said, 'You have got a blockage. I would suggest having this operation'. Well, that more or less made my mind up to go through with it, when I'd had that UDS . . . that was the final straw, yes. When I had that done, I decided then I was going to have it done.*
>
> *PS11, UDS*

> *It meant, meant that there's nothing else wrong with me, or any concerns . . . and that I just have to live with it really.*
>
> *PT6, UDS*

> *[UDS] gave them immediate information on exactly what the problem was.*
>
> *PT10, UDS*

Assessment results influenced treatment decision-making to varying extents. For some men, the assessments were essential to both their clinicians' and their own decision-making process (e.g. validating what was already suspected and/or providing a rationale for a treatment pathway):

> Interviewer: *Do you think the results of the assessments that you did helped make the decision for which treatment you'd get?*

> PT16, routine care: *Oh, yes, yes, yes . . . For both [clinician and patient], I think, because, I mean, although they weren't pleasant, you feel as though at least everybody's had a go, and they've done as much as they can. And I think when [consultant] sort of turned round and sort of said, 'Well, that confirms everything' . . . and he's not made any secrets of what he thought, so you knew what he was talking about.*

I didn't realise that I had a flow rate problem until I had this flow test done.

PT13, UDS

[The assessments] answered questions for the nurse and the consultant and then they could explain to me the implication of the results . . . if I hadn't been producing the flow that I was, you know, they said that would warrant a lot more investigation than the results that we have, which were quite satisfactory.

PT20, UDS

Others reported that, although assessments were helpful to clinicians, they had not personally found them helpful. Either because the participant was happy to defer interpretation of assessments and treatment decision-making, or because the participant already wanted surgery, but felt his consultant needed convincing evidence:

I guess it proved what I already knew. So I suppose it was for the doctor, I guess they needed more evidence than me saying my flow is not as strong as it used to be.

PT14, routine care

I probably had the idea that [surgery] was the correct way to go before [UDS], basically from what was happening to me, but then [UDS] validated it and therefore, you know, both the consultant and myself were on the same path then.

PT22, UDS

Of the 41 patients interviewed, five said that the assessments had not helped with the treatment decision. The most common reasons for this were the patient already knowing that he wanted surgery to alleviate his LUTS, the consultant being perceived as already decided on surgery, and the test results being inconclusive or not presenting any new information:

Researcher: *Do you think the results of your test, the test that you did, do you think those helped make the decision for surgery?*

PT3, routine care: *No. The results, not really. My symptoms, as I know them, convinced me that something needed to be done. Erm, the fact that we want to get on with our lives and get it sorted, hopefully. And, erm, we realised there was no other way, really, because the fact that it has gone on for 5 months, now, every day, the smell, and we're going to carry on like the rest of our days, might as well have something done to try and remove that.*

I mean nothing else was gonna work anyway, so it was a question of, did I put up with this, or do I have the operation. It was either have it or put up with the problems which would increase, and I wasn't prepared to do that.

PS9, UDS

A further three participants said they were unsure if assessment results had assisted with treatment decision-making, because either the consultant led the decision-making or their medication had not been changed:

Researcher: *Did you think that the results of the assessments helped you make the decision to take medication instead of surgery?*

PT2, UDS: *Well, I not make the decision myself the doctor, it's up to the doctor. If the doctor say 'You stay on the medicine', I stay on the medicine.*

Researcher: *And do you think that the results of the assessment then made the doctor decide that medication was better than surgery in this case?*

PT2, UDS: *Well, he had to make a decision and he said, 'Yes, you should take medicine'.*

Types of treatment decision-making

The type of treatment decision-making after LUTS assessment fell into one of three categories: (1) shared, (2) clinician led or (3) patient led. Overall, 20 out of 41 patients described the decision as shared with their consultant or clinical team, 14 patients described the decision as clinician led and seven patients described it as patient led (*Table 33*). Descriptions of clinician-led decision-making were more common among patients receiving conservative, non-surgical treatment for their LUTS, whereas patient-led decision-making was more common in patients who opted for surgery.

Shared decision-making was characterised by the patient and consultant or clinical team discussing assessment results and treatment options and together agreeing a course of treatment. Participants describing shared decision-making felt that clinicians were open and shared information and that they were involved in the decision:

No, I felt that, you know everything was, the information was well-shared, and I understand what I was doing, and yes, it was absolutely fine, and I didn't feel, you know, I felt well understood and that I was listened to and yes, I think they, they were very, very kind of open about whether I wanted to do the tests or whatever.

PT5, routine care

TABLE 33 Patient descriptions of LUTS treatment decision-making

Decision-making type	Described by:		Exemplifying quotation
	Pre-treatment patients, *N* (*n*, treatment decision)	Post-surgery patients (*n*)	
Doctor led	11 (8 conservative,[a] 3 surgery)	3	*I look on doctors like mechanics. They know best [laughs]. They fix my car; they can fix me* *PT11*
Shared	11 (5 conservative, 6 surgery)[b]	9	*Actually, they can't tell you, but they would probably recommend it to a certain extent, but you've got to make your own decision. But, yeah, they all supported me. Fantastic, to be honest* *PS5*
Patient led	3 (3 surgery)[c]	4	*You know, they give you all the information . . . but you still have to make that decision yourself and you just have to sit and . . . you have to sit on your own thinking about, you know, the consequences of what you're undertaking, you know . . . it's still my decision* *PT1*

a One patient said if his consultant had recommended surgery he would not have agreed to it as he did not consider his symptoms severe enough.
b One of the patients receiving conservative treatment had previously rejected the recommendation of surgery, one of the patients listed for surgery said he was still unsure, and another was having second thoughts and wanted to discuss with his surgeon an alternative technique ('UroLift'), publicised in the media and available privately.
c Although still listed for surgery, one patient had changed his mind at the time of the interview and the two others were unsure if they still wanted surgery.

Clinician-led decision-making was characterised by the clinician primarily making decisions about assessments and how to proceed with a patient's treatment and the patient abiding by this decision. Participants who described the decision as clinician led described not feeling involved in the decision-making, but most were happy to defer the decision to the clinician as the person best placed to make such decisions. In these cases, participants often did not feel that there was a decision to be made from their perspective, they deferred to the clinician's expertise and saw the treatment decision as wholly the clinician's decision:

> *The doctor is the expert . . . It depends on doctor, if doctor say I [should have] surgery, I have to go . . . it is the doctors, they had to make a decision, not me. If they said, 'well you need operation' then I have to make a decision, yes or no.*
>
> *PT2, UDS*

Some participants receiving conservative treatment and describing clinician-led decision-making did feel that they would have wanted more involvement and information if surgery had been recommended:

> *I mean, they decided that I didn't need an operation, whatever the operation was. He did explain it to me before, but because I didn't need it, I didn't take much notice after that . . . I wouldn't have surgery done, anyway. If they recommended it, I don't think I would have gone down that road. Don't fancy the idea of going into hospital and having, having an operation in, in that particular area.*
>
> *PT7, routine care*

Patient-led decision-making was evident when patients stressed that the treatment decision was theirs and that they had led and directed the process of reaching the decision, sometimes consulting family members. Some clinicians encouraged this approach:

> *The young – he were a young doctor – he explained what had happened, and then after he explained what I'd been doing, or how it had worked. And then he said to me, 'Go away, discuss it with your wife . . . You've got a 96% chance it does work, but you do get the odd 1 in a 100 that might, you know, it might not work, and you might have this problem, you might have that problem. Go away with your wife, discuss it, and then it's up to you whether you want it doing or not'. Which I did.*
>
> *PS11, UDS*

One patient described why directing the treatment decision was crucial for him:

> *The issue is in all of this, do you have enough information? Are you happy that the information is contextually correct and appropriate? Are you then happy that you've got the opportunity to make an informed decision? Yes. And there's no imposition by anybody . . . Well informed and then regardless of what happens, I am at ease with that situation . . . Being at ease with the decision is very important. 'Cause the alternative is to be uncertain: Have I made the right decision? What if this? What if that? And that sort of stuff is stressful.*
>
> *PT25, routine care*

Post-surgery participants' views of surgery

Post-surgery participants reported that their primary motivation for wanting surgery was hoping to eliminate or reduce the LUTS that they had been experiencing and the symptoms' impact on their lives. Participants wanted to return to a state that they deemed 'normal', in most cases referring to the way they were before the symptoms started:

> *I just wanted to get it sorted out, yeah. I couldn't carry on like I was, it was just . . . it was embarrassing and painful.*
>
> *PS11, UDS*

I got fed up of getting out of bed at least four times every night . . . It gradually got worse and worse and it worked out about six times a night on average . . . In the daytime, even not at work, I were in and out of the toilet a lot, most days of the week and I just got fed up with trying . . . it'd start and stop and have a job to start again.

PS16, routine care

I hoped that I was going to have to stop worrying about all the time when you go out, that you're sort of having to plan your routes around where there is a toilet . . . I just thought, well, we've got to go ahead with this, because it's no way I want to live with it, sort of, well the effect it was having on my lifestyle.

PS1, routine care

Participants also reported wanting to stop taking medication as motivation for surgery:

I was getting a bit sort of cheesed off with it suppose. I don't take pills a lot, I don't use them much at all, never really had paracetamol anything like that, don't have anything. So, to think about having to sort of take those all the time, and eventually when I did . . . go to the local pharmacy to collect my, you know, my tamsulosin and finasteride, I was just amazed at the amount of sort of boxes of stuff – then I felt Christ, I'm not going down that line, no don't normally take pills.

PS13, UDS

Expectations and experiences of bladder outlet obstruction surgery

Participants' expectations of their surgery varied across the 16 post-surgery participants and were influenced by the information they received from the urology team in advance. Some participants described not knowing what to expect in terms of the surgery and recovery, and felt that information provision had been poor:

Beforehand I wasn't really sure . . . wasn't really sure if it would work or quite what recovery would be like.

PS1, routine care

Well, I didn't know what it [surgery] were gonna be like.

PS16, routine care

Although some participants wanted more information, others felt that they didn't need a great deal and trusted that the doctors knew what they were doing:

Well I wasn't quite sure [what surgery would be like] . . . I tried to get out of him [the doctor] how on earth you go down through a tube and bore outside, but didn't get any answer to that I'm afraid . . . really just explained what was going to happen and it was the anaesthetic where the lower part of my body was completely dead for 6 hours . . . Didn't have a general, no. Spinal anaesthetic . . . It wasn't my choice, it's what they wanted to do. No [information], not really, no. You know, you hope people know what they're doing, they got on with the job, and it's all done and dusted.

PS14, UDS

Other participants felt that they had been given ample information and that their concerns or queries had been addressed, so that their experiences of surgery had met their expectations:

[My expectations for surgery and recovery were] exactly as it's going . . . we got a chance to do that [discuss queries] before, you know I saw [the surgeon] the morning of the surgery . . . as I say, the pamphlet that I was given detailed everything, and so, consequently, no real surprises.

PS9, UDS

Another way in which participants formed expectations of surgery was by talking to friends and family members who had the same or similar procedures in the past:

> *I thought it would be . . . I know a few people who have [had it] done, including a brother-in-law of mine, and I expected a lot of pain in the immediate aftermath and then a rapid . . . a rapid return to better . . . better urinary practice, you know, more regular, no leakage, that sort of thing.*
>
> *PS4, routine care*

A few participants reported that they had to wait longer than expected to have the surgery, which was difficult for them as they had to manage their symptoms in the meantime:

> *There was a fair waiting list. Initially I was down to be for surgery at the end of last year, would have been about this sort of time last year they said that, it was that but, I get the feeling when I sort of rang up and asked if there was any idea of sort of general type thing they said 'well no we don't really know', 'cause they had a lot of cancer type patients first, which obviously had priority, so ours, mine, obviously wasn't a priority one, as far as they were concerned. It was getting inconvenient for me . . . after that, it was just that, I think it was fairly . . . the sooner the better type of thing . . . You'd be in a queue which could be anything up to a year. So, thought oh no . . . That was a, I suppose that was a bit disheartening in that respect, having to wait more and more, because it was, it didn't give me confidence at all about sort of going out or that, [I] was forever, as I said, generally having to look for, trees and toilets and things.*
>
> *PS13, UDS*

Some participants described their expectations for the surgery being met or even surpassed by the experience of surgery itself:

> *I mean, really I thought it went much better than I was expecting, because I suppose it all went quite easily, and I knew, I sort of knew what to expect, so I wasn't too worried.*
>
> *PS1, routine care*

> Researcher: *Did your expectations of what was going to happen, of the surgery, match what actually happened? Was there anything that surprised you or . . . ?*

> PS13, UDS: *It was greater, it was brilliant, it was, no the whole thing was er, was a very good experience, very good experience . . . I actually . . . genuinely, thoroughly enjoyed the day . . . everybody, whether they were, you know the consultant or whether they were the anaesthetist or whether it was the cleaner on the floor, they were all brilliant and er, I totally enjoyed the whole experience.*

Participants typically stayed in hospital for 2 or 3 days when they had surgery, with some only staying one night. All expected to stay in for approximately 2 or 3 days, but those who were discharged earlier were happy to go home and did not feel concerned about this. Some participants had to return to hospital following complications, mainly on an outpatient basis, but a few were readmitted:

> *I was supposed to come out the Friday . . . but what happened, they took off the catheter – I had a bad time. They took off the catheter and then when I came home I couldn't pass water and so I had to go back [that night] . . . when I went home again I was having so [the next morning] I have to end up in – back in the hospital and spend a couple of days there.*
>
> *PS7, UDS*

Expectations and experiences of post-surgery recovery

Most participants reported feeling prepared for what to expect in recovery and having been given good information about recovery by the health-care team. Leaflets, the internet and other people who had been through similar procedures were the other main sources of information:

I don't remember exactly what I was told, but mostly, you know, I read it on the internet, erm, how long it takes and, erm, what sort of stay one should expect in hospital and most of the information, I get it from the internet.

PS15, routine care

Well, there were various booklets or papers that were given to me . . . And I realised that there was going to be problems over the first 6 weeks.

PS10, UDS

Many participants reported experiencing pain, discomfort and bleeding after surgery, and were expecting this. They knew that they would have to drink plenty of water and could develop an infection or need to have a catheter fitted:

If you're going to have surgery, you're going to have pain afterwards. That's probably a given to a certain extent.

PS5, routine care

However, some participants who reported these outcomes felt that they had not been prepared enough and would have liked more information:

When they sent me home that night, I went home equipped with a catheter which I had to use for the next 2 weeks. I hadn't expected that, but you know, I didn't feel betrayed or anything.

PS4, routine care

I suppose it does say in the pamphlet that you may get a urinary infection, but so maybe you know I should have been told more about that really, because I did suffer with it for a good week or so, thinking it was just normal rather than an infection.

PS2, routine care

I was bleeding quite a lot and it did carry on bleeding for about 2 weeks . . . That was a bit unexpected.

PS1, routine care

Some participants reported recovery being better than they expected and were surprised as to how little pain they experienced and how quickly their pre-surgery symptoms had subsided:

It was a lot better than I thought it was going to be, to be quite honest. Yeah. I recovered very quickly, which I found amazing, to be quite honest, but . . . Because you expect, when you have surgery you expect there to be pain for at least a week or so, don't you, a couple of weeks on and off. It comes and . . . It gets better day by day, like, anyway, but there was nothing particularly. You know, there was no pain or anything.

PS5, routine care

Recovery was instant. As soon as I went to the toilet, it was flowing fast. I could have a full bladder and it was empty within 20 seconds instead of 2 and a half minutes like it used to do. And it was brilliant.

PS11, UDS

For a few participants, recovery took longer than they thought it would; in some cases, this was due to complications following surgery:

After coming out of hospital, I started getting problems with my urine. Obviously, there was blood and you know that sort of thing, but, erm, I think after [2 weeks] I started getting a feeling of urine retention. And, erm, that night at about 10.30 p.m. it sort of completely stopped my urine, then I had to go to [hospital X's] A&E [accident and emergency] and I had a catheter put in. At the A&E I was given a rectal examination by the duty doctor – erm, I have written it down so I am reading it – who reported enlarged prostate, no hard stool and he prescribed an antibiotic called Nitrofurantoin.

PS15, routine care

Negative recovery experiences were made worse by a lack of consistency in health-care professionals, having to wait a long time to see a doctor after surgery and a lack of medical attention after being discharged, particularly out of hours. Some participants would have liked more opportunity to access specialist follow-up care. Having a telephone number for the hospital that patients could ring during recovery was valued:

The one thing I would say though is that I never saw the same doctor twice . . . So I saw the consultant, like I say, I saw the surgeon when I had the urine test, the flow test, then I saw him just before the operation, that same one, but after that, the one I saw was different, when they come round for like a bed check there's always a different doctor every time . . . Then when I did [see], you know, the urologist, the last one, that was a different urologist from the one who done the operation . . . so everyone was different and there was no continuity . . . It would be nice to have more continuity.

PS2, routine care

I think it [would] be much better if you could [contact] somebody for a bit of advice or even another backflush out of hours . . . I think a little bit more medical attention rather than trying to sort something out on your own [would] be better just for those initial 2 or 3 days when you have issues.

PS12, UDS

When they gave me that phone number when I went home and said I should call any time of night or day if I have any concerns at all, I think that had been pretty well covered and I didn't have to do that, but it was good to know that I could do that.

PS4, routine care

Some participants also reported being given advice regarding recovery, but feeling well and so not adhering to the advice. In some cases, this was fine, but other participants experienced negative reactions:

I think about 3 weeks and I was back to doing everything I, maybe not supposed to have done but I was doing it and it's never had any effect on me and, you know everything is working all right.

PS14, UDS

I went on holiday 6 weeks later, in fact after the operation . . . I climbed a big hill, which is 3800 steps and I came back down again and I found I bled the next day . . . I probably overdid it.

PS2, routine care

Symptoms post surgery

In general, participants were happy they had the surgery and would recommend it to others with similar symptoms. Some participants recognised the risk of symptoms returning and some were still experiencing

negative symptoms at the time of the interview, but, overall, participants reported that they had improved and that the pros of surgery outweighed the cons:

> *If I knew someone who had the same problem as me, I'd definitely advise them to get it done, because as I say, it has made a vast improvement so and there is that I'd need to do it again because it, the prostate can grow back . . . so I may have to have it again which wouldn't really bother me too much now.*
>
> *PS2, routine care*

Across the sample, the success of the surgery was categorised as high, moderate or low, based on participants' descriptions of their current symptoms and feelings. There was no apparent pattern between the success of surgery and whether or not participants had received UDS. For eight participants, the surgery had been highly successful, with a marked improvement in LUTS (e.g. improved flow, less urgency and complete emptying of the bladder), and high levels of satisfaction:

> *It's gone really. I don't have the flow problem. I don't have the leakage problem. I don't have to worry about whether I'm going to be caught out somewhere when I feel desperate need to urinate and I'm unable to do so in a civilised fashion. So, all of those concerns, I think have been addressed by the surgery.*
>
> *PS4, routine care*

For five participants, surgery had been moderately successful. These participants had seen their LUTS improve and were satisfied with the results, but still experienced some problems (e.g. urgency, retrograde ejaculation or slow flow rate).

> *It's better than it was. As I say, I can go to the toilet now and empty it out. Before I could go and probably, what, 10 minutes, quarter of an hour later – and same days it'd be back again. But it's improved.*
>
> *PS16, routine care*

> *Because I get retrograde ejaculations now . . . they said it was one of the side effects. So, I knew it would happen, so it didn't really matter . . . Although the medication I was taking was stopping ejaculations anyhow so . . . [it] didn't really make much difference really.*
>
> *PS2, routine care*

Three participants had seen no real improvement in their LUTS, or had seen improvements in some (e.g. nocturia, flow) but worsening in others (e.g. pain, incontinence). These three participants felt that their current symptoms were a direct result of the surgery or a side effect or indicated that the surgery had not been successful. Participants were told that their symptoms would ease over time and wondered if they were still recovering between 48 and 63 days after surgery. Participants were aware of the risk of side effects, but did not fully understand the symptoms they continued to have and would have liked more information:

> *The symptoms haven't really changed. I'm still getting the same feeling of pressure and um frequency um, still getting those, but again I'm to expect that . . . Because I've got to get the muscles in that area back to how – they've been messed around for 2 years and I'm not going to get them back in 2 months.*
>
> *PS9, UDS*

> *I may have ongoing infection, poor bladder emptying, or renal obstruction, all of which need to be assessed. [The consultant] recommended urgent urology review . . . After that, I have met [surgeon] and he has assured me that, erm, the operation went very well and the symptoms which I have . . . like tummy swelling or [pain in] my right side waist, they are not connected with the surgery. They may be something else, you know, which is resulting in this.*
>
> *PS15, routine care*

Participants discussed having to adjust to their new situation, whether this was not having symptoms they had got used to, or having new ones because of the surgery. They used medication, self-catheterisation or incontinence pads to manage any ongoing symptoms:

> *Basically, if it gets bad I might take sort of ibuprofen or paracetamol but basically you soon get used to it and you just adapt to it, well I've adapted to it.*
>
> *PS6, routine care*

> *The only drawback is having a self-catheter. But I've got used to it now.*
>
> *PS8, routine care*

Clinicians' views and experiences

Methods

Trial design
In-depth semistructured interviews were conducted with clinicians involved in UPSTREAM.[19]

Recruitment and sampling
Clinicians were purposefully sampled in relation to (1) trial centre and (2) professional role. Urology consultants were recruited via an e-mail invitation from the trial chief investigator. Research nurses were recruited via an e-mail from the UPSTREAM lead research nurse. E-mail invitations included the participant information sheet and the researcher's contact details. Consultants and research nurses who expressed interest in participating were contacted by a researcher to arrange a convenient time for an interview.

Interviews
All the clinician interviews were conducted by telephone. Interviews followed topic guides developed at the start of the trial based on the literature, the qualitative trial aims and input from the TMG (see *Appendix 6*). They were also piloted and refined prior to data collection. The topic guides were further refined as data collection progressed, in response to emerging findings. Consent was taken verbally before the interviews began. All interviews were digitally audio-recorded, professionally transcribed verbatim, anonymised and checked for accuracy prior to analysis.

Analysis of the interviews
Analysis followed the same process as for the patient interviews (see *UPSTREAM participants' views and experiences*, *Analysis of the interviews*).

Results

Participants
A total of 21 interviews were conducted [15 consultant urologists, five nurses (four urology nurses and one research nurse) and one UDS technician], with staff from 18 different trial centres. Number of years' experience in their role ranged from 3 to 37 years (mean 16 years). Fourteen of the 21 clinicians had previous experience of using UDS (67%). Interviews were conducted between July 2016 and March 2017 and had been involved in the trial between 12 and 34 months after trial initiation (mean 23.4 months, SD 5.53 months).

Clinicians' views and experiences of UPSTREAM
The consultants interviewed were responsible for championing the trial in their centres, with some being trial co-applicants. Most were involved in identification of patients and trial recruitment. Some were involved in the UDS testing procedures and all were involved in the interpretation and discussion of the results with patients, as well as decision-making concerning surgery. The nurses interviewed were mainly

involved with inviting patients to have the test and performing the test. A few of the nurses were involved in the interpretation of test results, but none were involved in treatment decision-making.

Clinicians expressed mostly positive views of the trial and enjoyed being part of it. They felt that it was an important and timely trial. They also welcomed answers that would standardise pathways and decision-making in current inconsistent systems:

> *Those practised by different consultants in different departments are so diverse . . . so it's hard to know how consistent it is.*
>
> <div align="right">*Clinician 11*</div>

Clinicians reported having good support from the trial team that helped facilitate the trial at their centres. Updates on recruitment and the progress of the trial and reminders were welcomed by clinicians, and they liked the proactive nature of the trial team:

> *Support from [trial team] has been very good. I think chasing up reminders has been very good. I think the meetings organised to give updates, I think that's been very good.*
>
> <div align="right">*Clinician 15*</div>

Some clinicians voiced concerns over the inclusion criteria, leading to heterogeneity of trial participants, and felt that this may lead to difficulty in interpreting and translating the trial results to inform practice:

> *You know, it's an interesting question because it's a test which is not very pleasant, so I will see what it will tell us. I think we'll end up being told a group being defined in which it's helpful. I don't know that there'll be many surprises and I wonder if it'll be very useful clinically, 'cause I suspect you'll have to remember quite a few features about that group and you may not knowingly meet that particular group. It might be difficult to define the group that apparently UDS is helpful in.*
>
> <div align="right">*Clinician 4*</div>

Although clinicians had varied views on the usefulness of the trial outcomes, all believed themselves to be in equipoise for whether UDS or no UDS would be more effective:

> *On the whole them seemed . . . I think relatively balanced with equipoise and make the call as saying that we don't really know exactly who should or should not have tests.*
>
> <div align="right">*Clinician 2*</div>

Identification and recruitment of trial participants

Most clinicians felt that their centre had done well with recruitment, but some issues were reported when identifying and recruiting patients, including clinician engagement with the trial strain on clinicians' time in already busy clinics, some patients' reluctance to take part and variance in care across clinics.

Some clinicians reported resistance to recruiting to UPSTREAM from hospital management and colleagues who had concerns around the pressures on existing waiting lists and patient numbers. They did not want to add more patients to already busy clinic lists as a result of the trial. It was also difficult to identify and recruit patients when other clinicians were not pro-research:

> *To try and push in study which would involve doing UDS in half the patients that we recruit, it was a bit tricky when it came to management, when it came to other clinicians who were, who were a little bit concerned about what would happen to the waiting list and the numbers. So if there was a way by which we could have expressed or had additional sessions to do the UDS of those patients who, who were selected in the intervention group then maybe, I don't know, maybe it would have made a difference because then the other clinicians would have probably tried to come on board by offering, offering their patients for the study. But I think that was probably something that played, well, not*

made terribly obvious but it was there going around in the background that, that there might be an additional workload on the UDS list.

Clinician 10

It's getting others who see patients . . . but don't have the same level of interest that I or one or two others do. In other words, I think patients are quite interested and pro-active in the trial. I think it becomes a lot harder selling it to colleagues continuously when they don't have a direct interest in the trial.

Clinician 15

Diversity and consistency of care across centres were reported to be challenging when recruiting patients into the trial. Owing to having several consultants and clinics spread out across centres, research nurses could not always access all patients in all clinics. This was further complicated by patients not always continually seeing the same health-care professional:

Just identifying the patients, once we knew about the patient we could recruit, but the patients present at different consultants' different clinics, different community clinics so trying to find out, identification is a difficulty.

Clinician 11

Some clinicians reported patients being reluctant to take part in the trial for a range of reasons. Clinicians reported some busy, younger patients not wanting to attend additional hospital appointments; patients wanting to progress with the most effective treatment available and not wanting to risk being randomised to another assessment; and the invasive nature of UDS potentially putting patients off:

We still struggle with recruiting. We screened a lot of patients, but we haven't recruited that many and I think it's mainly because patients are reluctant to undergo these procedures. In fact, there is certainly reluctance to undergo UDS, that arm.

Clinician 7

So it was difficult to enrol young patients because it's an extra test and they need to come once more to the hospital and I can understand it's not very pleasant.

Clinician 5

A key facilitator to trial recruitment and success at centres was having a research nurse who had dedicated time to identify patients, explain the trial and recruit them to the trial. Dedication and available time was important to clinicians, but the presence of a research nurse was particularly helpful and clinicians felt that they could not have recruited sufficiently without them:

I think the most important thing is dedication and time, dedication from our point of view as a clinician to identify and initially explain the trial, remember you're identifying these patients in a busy clinic and you then can't afford to get them back to another clinic to go through it again . . . so again it's having the nurse with the time because I effectively patient identify and then I'll introduce them to the trial to see whether they are interested, once I've got their interest I then get the research nurse to sit down with them half an hour, giving them all the information . . . So I think the time spent by the research nurse, that is crucial.

Clinician 15

I think, as I say, we're lucky because we have a research nurse who has basically has taken on all the administrative side of it, all the contact with the patient. So our involvement has been to carry on and do the procedures.

Nurse 4

Impact of trial participation on usual practice

There were mixed views on if and how participating in the trial had affected usual practice and interviewee workload. Some clinicians believed that the trial created an artificial decision process: if the trial was not in place, patients would not perceive a choice in whether or not to have UDS, but rather take it as the next step in investigative tests as advised by the clinician. This, therefore, changed the way that they perceived the test from a needed test to an optional one:

> *I think if we, outside the research, if you were to say to patients 'you need another test and this is what it would involve', I, we used to sort of put it across in a way that this is the next step, but in the research model when we were recruiting them and giving them the details of what could happen, that they may or may not have UDS, it's a slightly, slightly different, different scenario because what, they knew that this is probably an optional test rather than when you really see a patient and you think like, 'OK, I need urodynamics here', you tell the patient that you need UDS so the patient gets the impression 'OK, if I don't have this test we don't know what needs to be done next'. So that difference was there.*
>
> *Clinician 10*

> *Well the trial has changed us because we don't have to make that judgement whether we, yes because we have to, we look at someone at say you know, we enrol them for the trial and then you see them with their UDS at a part where you might not have already have done them or you have made a decision to do surgery in the absence . . . So it has changed our usual practice.*
>
> *Clinician 4*

In some centres, the trial created extra steps in the patient pathway. There was extra diagnostic testing in some patients who would not otherwise have had UDS, which was perceived to delay the clinical decision-making. Also, an additional follow-up consultation was necessary at some centres to discuss treatment and management post assessment:

> *It's got an impact in as much as it's another step in the diagnostic pathway, so the patient that was going to have surgery has [had to have UDS] in the study but wouldn't have had it clinically, then it takes longer to get that clinical decision.*
>
> *Clinician 11*

> *After the UDS I would normally discharge a patient if there was nothing to be done. But with the trial there is a further follow-up consultation and that has added to the outpatient activity.*
>
> *Clinician 14*

Some clinicians interviewed found it difficult to fit the additional work created by the trial into the time they currently had available in practice. One clinician reported using their allocated administration time to run additional clinics:

> *The challenge was, if you're going to recruit a person, it takes that additional bit of time and generally most of the clinics that we do are overbooked, so the situation is such that a clinician who's trying to do their clinic, his main aim is to try and finish doing what he needs to do or what she needs to do within that, within that space of time. Generally, it's not always efficient because the clinics always finish off late and going through the extra bits there is no help, can be, can be difficult who is not doing it on a regular basis . . . I don't think we were able to overcome it.*
>
> *Clinician 10*

> *Well, a lot of extra work . . . overbooking my clinic in order to fit these [trial] patients in, I have to open an extra clinic . . . I have to either book the list or I have to create an extra list in my time which is designated for my admin[istration].*
>
> *Clinician 12*

Some clinicians struggled for space in the hospital to perform UDS as part of the trial:

It hasn't particularly added to the workload where we were going to see someone anyway . . . we're just short of capacity in a proper area to put them in, which made life a little difficult.

Clinician 2

Not all clinicians reported adverse impact on their current practice, with clinicians who already performed UDS routinely easily able to absorb UPSTREAM patients into their usual practice:

It's slotted into our standard practice quite well in the sense that the UDS clinics that I already do, I've included the men in those without too much trouble, so it's not put undue pressure on our service. The other specialist clinics that I do, the non-UDS patients have been fitted in to that, and, again, there hasn't been a particular problem with service impact.

Clinician 13

It's no different to what we used to do when we weren't sure if someone should have a TURP, we did UDS. They came back for the results and we tell them we think they need an operation or that it wouldn't help them. So no change.

Clinician 4

Some centres completed all tests and discussed the results in one session, a 'one-stop shop' (Clinician 5), whereas, in usual practice, patients would have to wait for a follow-up appointment, which would have resulted in a longer process:

But if you look at it, that process [usual practice] takes longer . . . Because you're talking about two, three times of clinic. One, the patient comes in, two, they refer the patient to the UDS, so he or she, well, we're talking about 'he's now, will receive an appointment letter . . . They come in . . . have to review the results and contact the patient, or get another appointment for this patient to come and discuss the results. Whilst within UPSTREAM, we do all these things on the same day, within the same appointment.

Nurse 1

Some clinicians reported increasing the use of UDS outside the trial in their usual practice as a result of seeing the benefits of UDS in the trial:

Yes, it has [affected practice,] it has and I, I have found that I tend to offer urodynamics more to the group of patients that probably go through me now.

Clinician 10

Clinician views of urodynamics

Views on UDS varied across clinicians, as did the extent of UDS use in usual practice before the trial. Most clinicians spoke positively about UDS, with some saying that it was underused, but a few expressed more negative views and believed that UDS was overused as a result of a lack of evidence on when to use it:

I think it's [UDS] important to investigating people with complex urinary tract symptoms and I think it's something which is really necessary in a large number of people. I also think it's underused.

Clinician 7

It's [UDS] requested for reasons that aren't clearly thought out, so potentially it could be being overused.

Clinician 3

Some clinicians who expressed more positive views used UDS more routinely before surgery, with one clinician reporting using it before any invasive therapy:

> *I love it, it's like lifting the bonnet on the bladder, so I always like seeing the outcomes, but have some colleagues who would always do UDS before a TURP and I wouldn't.*
>
> *Clinician 4*

> *I think it's very useful . . . I think essentially that at the end of the day, if you're going to perform an invasive procedure, i.e. an operation, then you've got to be sure that's the right operation, then you should definitely do it . . . personally I think it has a big role.*
>
> *Clinician 9*

> *My personal view is that the vast majority, if not all those should undergo UDS before TURP to confirm BOO.*
>
> *Clinician 13*

Some felt that the usefulness of UDS was limited and the test was not always accurate because of the artificial nature of UDS:

> *I like it. I don't do it. I like it. It gives a slight objectivity, but that's all.*
>
> *Clinician 14*

> *Occasionally, you don't get an answer to your question . . . you're not always able to reproduce the symptoms that you're trying to reproduce in an artificial clinical setting, so it's not always diagnostic.*
>
> *Clinician 3*

Clinicians who did not routinely use UDS before surgery usually performed UDS after all other tests had been done and when results had not produced a clear answer as to whether or not surgery should be performed or when the diagnosis was unclear:

> *If the patient has a history of enlarged prostate, poor flow, I tend to carry straight on with surgery without UDS.*
>
> *Clinician 11*

> *It's the last test that's being requested . . . just to confirm their findings . . . to formulate a correct, the best management for the patients.*
>
> *Nurse 5*

There was consistency across clinicians when describing patient symptoms that would warrant UDS. These included the patient being considered young (aged < 60 years) or old (aged > 80 years), having chronic retention, bladder underactivity or complicating factors such as neurological conditions. UDS would usually be used in patients for whom medication or conservative treatment had failed or who had previously had surgery:

> *Large residuals, so chronic retention. If the flow rate is variable or borderline . . . there was a mixture of storage and voiding symptoms . . . but typically after they have tried medical therapy, and usually failed or not had sufficient benefit.*
>
> *Clinician 2*

> *The typical indication would be extremes of age, i.e. young patients and older patients and also those who specifically failed after previous intervention or surgery.*
>
> *Clinician 8*

Clinicians reported different practices when it came to who performed the UDS test and interpretation of test results. Most UDS clinics were nurse led, with some mixed nurse and consultant led, and a minority being consultant only. In most centres, consultants interpreted the results and discussed the findings with patients, but some had nurses deliver the results and if needed, scheduled a later discussion about treatment with the clinician:

> So the specialist nurse who is doing the dynamic test and is talking to the consultant and then he's [nurse] delivering the result as in verbal and some written information. And the exacting of that results, so if the patient needs something to be done accordingly. But then we have another proper discussion and the clinical discussion with the patient and the consultant.
>
> *Clinician 5*

> I usually do it [explain UDS results] at the time of the test. We are there and we explain what we found during the urodynamic test ourselves.
>
> *Clinician 7*

Clinician perspectives on the role of urodynamics in lower urinary tract symptoms treatment decision-making

Clinicians reported using the results of UDS testing to make the best decision possible for patients. The results helped clinicians to decide when not to go ahead with surgery and, if it was appropriate to perform surgery, the risks and benefits of surgery based on patient characteristics. In addition, UDS testing allowed clinicians to assess the potential success of surgery and if there would be any foreseeable issues or complications if surgery was performed. Clinicians also believed that the test helped the patients with improved understanding of their symptoms and treatment options, including surgery:

> If properly indicated it really does affect our decision and help make the better decisions for the patient.
>
> *Clinician 8*

> I think yes, it has a positive, positive impact for the patients and the doctor and it can help inform subsequent management and also help the patient understand in a bit more detail about why they've got this urinary symptom.
>
> *Clinician 1*

Clinicians reported that the UDS results provided clear-cut information about what was happening with the bladder and prostate (e.g. if there was a blockage). This helped them decide whether or not surgery would be suitable and provided a rationale for their decision to explain to patients:

> I think the advantage is that is gives good or accurate information about what exactly is happening with the bladder or the prostate. And it helps surgeons make decisions on whether to operate on someone or not.
>
> *Nurse 1*

> You're going in with the strength of knowing the results, i.e. you are definitely obstructed you're definitely not obstructed, you're definitely stable, no you're unstable. So you have the strength of the test behind you.
>
> *Clinician 15*

Urodynamics provided a way of assessing the risks against the benefits of performing surgery, in order to make a final treatment decision. This was particularly important in those patients who may be unfit for surgery or for whom surgery might present additional risks:

> So if I've got particular concern about the risks of the procedure, and I want to be sure that the potential benefit or it's so that they might not necessarily benefit but in order to make that kind of

judgement, if I want to be able to tell them [the patient] 'you know I think you've got a much more better chance of success based on your UDS and therefore we might go ahead'.

Clinician 4

I think UDS is beneficial in some cases where the picture is not clear cut so you cannot definitely say that OK, surgery is going make an improvement.

Clinician 10

Potentially avoiding an unnecessary operation, exposing them to the risk without the benefits.

Clinician 13

Urodynamics also helped clinicians understand what the outcomes of the surgery may be and whether or not any complications may ensue. Clinicians believed that the UDS results helped to facilitate a more informed conversation with patients about their symptoms and treatment options. The results could help patients understand the need to go ahead with surgery or have pharmacological treatment, and what the complications may be:

It might show that a very overactive bladder or small capacity bladder and that might make you a bit suspicious about any urgency experience after surgery.

Clinician 1

I think patients should have it [UDS] done because we have had patients . . . who had prostate surgery and actually they have been wetter after surgery than beforehand, because they hadn't had UDS studies in the first place . . . I think it helps for the patient, because to be worse after a procedure is not a good idea.

Nurse 3

So the situations where the bladder is underactive, they're not able to empty their bladders completely. I mean, even in those cases where if we were to offer them outflow obstructive surgery we would be able to explain to them, 'see your bladder is underactive, you may not always be able to empty your bladder completely', because some of them feel that it is surgery that is going to correct all the symptoms and it helps to explain to them some of the things which has less likelihood of getting corrected.

Clinician 10

Although all clinicians reported that they used UDS to inform treatment decision-making and facilitate their conversations with patients about these decisions, clinicians described various ways they did so. Some described informing the patient of the test results and providing their view on whether or not surgery should be performed. Some provided the test results and let the patient decide whether or not to have surgery, and others discussed results and options in a shared decision-making process. Some clinicians felt that the tests were more for the benefit of the clinical team and its decision-making rather than for the patients:

I tailor the information that I give according to what they want, what they understand, but, in general, they are making an informed decision . . . I don't usually recommend [surgery] unless they specifically ask me to. I just tell them it's their decision and I give them the information.

Clinician 3

They [patients] rely on my summation of all the tests they've had done. They are generally influenced by your decision, probably the odd one or two patients that will ask very critically every bit, the majority won't.

Clinician 15

The patients share that information and they have a say in the decision-making process. If we can give a bit more information to the patient about the probability of success based on UDS, they can make a more informed decision.

Clinician 1

I don't think the patient is that bothered. It's the symptoms that are important. What the test shows is not that important to the patient. I think the assessments play a role for the clinician [in decision-making] but whether they play a role for the patients . . . I'm not sure.

Clinician 7

Views on potential urodynamics roll-out

Clinicians agreed that UDS should be implemented more widely if the trial showed a strong indication for patient benefit. They stressed that a good evidence base would be needed to convince clinicians to use UDS more widely. Clinicians would want clear indication, protocols or instructions for who to target and which patients would benefit most from UDS (e.g. age, existing conditions):

I think the question is not just that studies show it's better, it's quantifying that benefit and how much benefit that it can have in patients. Had they had surgery when they shouldn't have done or were appropriately denied surgery. So, you know, it's not just the benefit, it's more the degree of benefit needed . . . maybe that would be an important parameter to know.

Clinician 1

I think you'd have to have fairly convincing data as it is required.

Clinician 14

I mean, if the trial comes up with something that provides good evidence of benefit and it's cost-effective, then I would hope that people would take it up . . . it's an invasive test which has got side effects and it's not particularly pleasant test for people to have. And it costs money and takes time . . . there's a risk associated with it so there has to be benefit.

Clinician 3

Clinicians discussed what would be needed to get wider UDS use into practice, should it be found to be effective. Clinicians believed that the impact of rolling out UDS more widely across the NHS would depend on whether or not centres were currently using UDS, the extent of resources available and the interest and engagement of clinicians:

It depends on your centre, I imagine somewhere . . . who uses them [UDS] a lot, it might not impact on the resource because they already have a [team] that's keen on the UDS. It's centres like our own, we're probably somewhere in the middle and then there may be centres that don't use UDS very much at all who may be having to create a whole new service to support that, or a centre that doesn't have urodynamics on site that's going to have to refer their patients to another organisation, so it will be dependent on the organisation and on what the additional resource will be.

Clinician 11

I think the major problem is that if people aren't doing UDS, then obviously you have to set up the service, which is always going to be a problem.

Clinician 6

I think we have the capacity. We have the staff, we have the kit . . . suppose you would need an extra UDS list . . . I think you'd need an extra list . . . that's an extra one a week.

Clinician 4

Some clinicians stressed the need for consultants to be involved in interpretation of the more complicated UDS test results. However, they recognised that not all clinicians have an interest in UDS and some may prefer to pass the tests to nurses to conduct and interpret. Therefore, if UDS were to be introduced as standard practice, clinicians would need encouragement to use UDS. National guidelines may help facilitate this engagement:

> *I think what is happening at the present time is that UDS is not something that any clinician willingly wants to take on, so it's not a topic which excites most clinicians . . . so I think encouraging more clinicians to be involved would be the next challenge because it's . . . very easy to say, 'OK, I've got a couple of nurses that I'm going to train and ask them to do UDS'. I mean some of the straight forward UDS results that do get produced by this test are straightforward but, then you do get certain tests which won't ultimately answer your question and having a clinician is necessary to make sense out of it.*
>
> *Clinician 10*

> *A lot of it would come down to whether or not the surgeons bought into the fact that it was, erm required . . . I think it comes down to if there was some kind of . . . some sort of NICE guidelines that all patients coming in for TURP must have UDS as part of their pathway, then it becomes something that we have to work out how to do, as opposed to that we have to decide whether we want to do it.*
>
> *Nurse 4*

Clinicians felt that additional staff training would be required, which would be difficult owing to current problems with the availability of training courses and where the courses were run:

> *Staffing, because you need technicians who can operate the machine. You need to be trained how to do it . . . the training of urology staff on how to perform UDS and, you know, how to capture all the relevant information should be advocated for, especially in hospitals where it is being performed . . . but if they have a machine, and it's being performed, then I'll recommend that staff are trained on it. The only problem is different hospitals who have different machines, which, you know, operate differently. But wherever you are, wherever you find yourself, get yourself trained.*
>
> *Nurse 1*

> *Staff training would be a concern. Like when there's only two of us was doing it, and not all hospitals, I think, have got their own UDS unit . . . it's not a regular, the course itself is not a rolling throughout the year, it's a specific day, a date I think, it's in a year and that . . . the venue is really far. In [place] we were lucky that when we were to have the UDS, you have attend the UDS course, it's [local] and that's accessible to us, otherwise you have to go all the way to [place].*
>
> *Nurse 5*

Clinicians also anticipated logistical issues with space availability; at some centres, clinicians reported that there would be no rooms available to conduct UDS:

> *Well, you need to find first of all the right resources . . . particularly in our hospital, we're very short of available rooms, so even though we would like to do more UDS, we can't because there's no other times when we can actually have a room available.*
>
> *Clinician 12*

Clinicians highlighted the strains that some current services were already under and voiced concerns that routinely introducing UDS would have a detrimental effect on services:

> *The main thing is the increased capacity you would have to have . . . I think there would be a big challenge in the NHS . . . a lot of units do have problems trying to provide UDS support for the team.*
>
> *Clinician 9*

We know our nurses are struggling to cope with the workload we've got and the space we've got which is very little for 10 consultants and others. So I think trying to routinely [conduct UDS] . . . would just break the service currently, or we'd have to give something else up.

Clinician 2

Clinicians were also aware of the cost implications of additional testing and said that they would need a clear indication of the NHS tariff for the test and how it would be paid for:

The other thing is . . . where do UDS fit in to the NHS tariffs, you know, with the funding stream and how UDS is funded within the NHS. Are you expecting the costs [to be accommodated], or do you get extra funding to carry out UDS?

Clinician 1

Even though additional resources would be required, clinicians felt that, if the evidence indicated that UDS should be used and would benefit patients, then it would be worth it:

There's implications, but if high-quality research suggests that the use of UDS guides us towards the right patients who are likely to benefit the most . . . high-quality health care that's cost-effective as well, it would reduce the risk of operating on people who will get no or minimal benefit and, hopefully, those that have surgery will be satisfied with the treatment and will be the right group to operate on; so yeah, it would be worth the additional resource if it was needed.

Clinician 11

I think we'll do that [increase capacity] because if it brings benefit to the patients, we will be happy to do that.

Clinician 5

Discussion

Summary of findings and implications

UPSTREAM was viewed positively by both trial participants and clinicians, with participants valuing the research for ensuring best practice and clinicians reporting being well supported by the trial team. Most patients who declined trial participation reported a preference for either receiving or not receiving UDS, or a treatment preference. Future studies can overcome this issue by encouraging recruitment staff to explore arm or treatment preferences during recruitment and, by doing so, help potential trial participants to express their concerns, discuss equipoise and randomisation, and reach an informed decision about trial participation.[55] Trial arm preference was also expressed by some of the trial participants interviewed, particularly those men in the UDS arm, who felt that they may receive quicker and better care and treatments, would know more about their condition as a result of UDS, or that it would help with decision-making, which overcame concerns about the procedure.

Most participants reported that UDS testing was acceptable, despite a minority of participants experiencing pain, a UTI or embarrassment.[48] Participants valued UDS for its perceived accuracy and the information it provided about symptom aetiology, and thought that it gave more insight into LUTS than other tests. Reflecting clinical guidelines,[4] UDS was carried out after other assessments, such as uroflowmetry and completion of bladder diaries, and was therefore perceived as completing the possible assessments comprehensively. For some men, the results of UDS played an important role in treatment decision-making.

Levels of discomfort and embarrassment experienced during the test varied, but all the men were prepared to have the test again, if needed, despite any negative experiences. Embarrassment was related to the intimate nature of the test and a lack of awareness or preparation for its effects. The keys to minimising embarrassment were good communication and privacy; ensuring men knew what to expect; limiting the number of staff present; and introducing staff and explaining their role. For some men, the sex

of the person performing the test was important. Information deficits were reported before, during and after the test, and there was variability in how and when results were explained, and the adequacy of explanations.

These findings have clear clinical implications. It is essential that clinicians inform patients in advance about what to expect during and after the test, including the risk and treatment of UTIs. Patients should be informed that passing urine during the test can be associated with spraying and that this will easily be dealt with afterwards. During UDS practice, efforts should be made to limit the number of staff present in the assessment room and ensure that maximum possible privacy is maintained. Patients should be introduced to the clinicians present and informed of their role, with agreement sought for the involvement of trainees in the procedure. After the test, patients should be allowed to get dressed in their normal clothes before the concluding discussions. Clinicians should discuss side effects with patients and what to do if they experience any problems. Clinicians and patients would ideally discuss the results of UDS testing on the same day as the test or shortly after, with the detail and depth of the explanation in line with patients' preference.

Our finding of variability in the degree of privacy, dignity, discomfort and information provision experienced by patients undergoing UDS testing indicate that staff require training and guidance in these areas.[48] Current UDS guidelines[4] omit guidance on how to ensure that urological assessment is patient centred, yet this is crucial to ensure positive experiences among patients. Inadequacies in UK training in the conduct of UDS investigations have been reported previously;[56] our findings suggest that such training needs to include the sensitive conduct of UDS testing and associated information provision. Our recommendations for best practice provide[48] support for the ICS statement that patients must receive proper information in advance of testing,[57] and our findings inform the medical community about the appropriate content.

Most trial participants who received UDS felt that it was useful in treatment decision-making: UDS could help clarify why they were experiencing symptoms, validate what they and the clinician had suspected, provide reassurance, demonstrate that they had a problem that needed treatment, help them understand what treatment was required and provide a conclusive answer that there was an indication to undergo surgery. Shared decision-making was the most common approach to treatment decision-making described by participants, followed by doctor- and then patient-led decision-making. Although some patients felt that they had the final choice over treatment decisions, others felt that they were not involved in treatment decision-making and UDS results were for the clinician to use in making a decision. Shared or patient-led decision-making was reported more among men who had received, or were scheduled to have, surgery.

The main motivation for surgery reported by trial participants was the desire to reduce or eliminate their LUTS and the impact of LUTS. Expectations for surgery and whether or not these were met varied between participants, and ranged from not knowing what to expect to expectations being surpassed in some cases. Information provision regarding surgery was also reported as varied in terms of what was provided, but also what participants wanted. Some participants reported having enough information but others not enough, some patients preferred not to have extremely detailed information whereas others wanted more detail. An unexpected lengthy waiting time for surgery was problematic, as participants had to continue to live with unwanted symptoms indefinitely and more clarity about what to expect in terms of accurate waiting times would have been valued.

Most trial participants reported being prepared for what to expect with post-surgery recovery and these expectations were mostly met. Again, there was a range of reported experiences of recovery and some participants experienced unexpected side effects or complications. A few participants did report negative experiences that might have been avoided with continuity of care and facilitated specialist aftercare. Participants felt that information about adhering to recovery activity could have been clearer, particularly in those who felt well and who engaged in work or other activities, which resulted in negative consequences.

Overall, participants felt that the decision to go ahead with surgery was the right one: 13 out of 16 post-surgery participants described highly or moderately successful outcomes and almost all patients reported the pros of surgery to outweigh the cons. However, several participants described having to adjust to their new situation, whether this was improved symptoms or negative side effects.

Clinician views on the acceptability and usefulness of UDS were varied, as was their use in routine practice. Most clinicians viewed UDS positively and believed it to be underused. These clinicians were more likely to use UDS routinely before surgery. Some clinicians viewed UDS more negatively, felt that it was currently unclear when UDS should be used and therefore felt that it was sometimes overused in usual practice. Some felt that the test's usefulness was limited, as it was conducted in an artificial setting. Variability in practice across centres resulted in heterogeneity and inconsistency in who performed the test and how and when test results were discussed with patients.

Clinicians valued UDS when making decisions about treatment and condition management, particularly in judging the appropriateness, probable benefit and potential risks of surgery. UDS provided additional information to supplement other urological tests, which helped clarify the patient's situation and helped when discussing possible treatment options with patients. Clinicians used UDS results in different ways, with different types of decision-making being reported. Some clinicians felt that the test results were primarily for them to make a decision, which they then advised the patient of, whereas others used this information to facilitate a shared decision-making discussion with the patient as to how to proceed.

Clinicians agreed that UDS should be implemented more widely in routine practice if there was a sound evidence base to do so. They recommended clear guidelines to enable clinicians to implement UDS appropriately. The impact of implementing UDS more widely would depend on existing practice, resources and clinician interest. Levels and availability of training in UDS would need to be considered and addressed, as would staff capacity and logistical issues, such as having appropriate space to conduct the test.

Main findings in the context of the wider literature

Previous studies have utilised qualitative methods to investigate patients' views, experiences and health beliefs about LUTS,[58–61] triggers for and barriers to help-seeking,[62] and perspectives on treatment outcomes.[63]

Existing evidence regarding the acceptability of UDS is largely from questionnaire-based studies and suggests that UDS is well tolerated,[64,65] but can also be uncomfortable[66] and might cause complications,[67] although the extent of these is unclear.[68] To our knowledge, UPSTREAM is the first in-depth study of men's attitudes to and experiences of UDS for LUTS, using formal qualitative research methodology.[48]

Our findings support those of Scarpero et al.,[69] who found that UDS was well tolerated, with patients experiencing minimal to moderate discomfort and embarrassment. As in Shaw et al.'s [70] mixed-sex (female-predominant) qualitative study, embarrassment was related to the intimacy of the procedure and privacy was therefore important. Previous research has suggested that UDS might be less acceptable to younger patients.[71] This was not evidenced among our participants, but this could be due to the older age group of men with LUTS in our trial sample (age range 52–89 years). Yiou et al.[71] found that being aged < 54 years was sometimes associated with experiencing painful sensation during UDS in a cohort of 68 men and 103 women.

Little has been published about clinician views on UDS, with existing literature focusing on female urinary incontinence and the use of UDS before surgery.[72] Our findings are consistent with results reported from a survey and interviews,[72] which indicate polarised views in clinicians on the usefulness and current use of UDS. Hilton et al.[72] found that the majority of clinicians used UDS before surgery, with some reporting UDS as essential in adding to the overall picture, to help inform the best course of action and to provide a safety net for reducing the risk of inappropriate surgery. Some clinicians, however, reported limited use due to a lack of clarity in existing guidelines on how and when to use UDS.

Strengths and limitations

The strengths of the current study include our focus on an in-depth understanding of patients' and clinicians' perspectives on, and experiences of, the trial, UDS and surgery; the recruitment of a large and diverse sample of trial participants in terms of age, symptom burden and treatment decision; our understanding of clinicians in terms of centre and professional role; and the attainment of data saturation. Not all the trial participants included underwent UDS, as the views and attitudes of those who have not had the test are also important in considering acceptability in the target population. Analysis demonstrated a high degree of similarity between pre-treatment and post-surgery interviews, with both containing positive and negative views and experiences of UDS testing, not dependent on treatment decisions or surgery outcome. However, most patients interviewed had consented to a trial in which there was a 50% chance of randomisation to receive UDS, so those totally opposed to it might not have consented to the trial; this should be taken into account in interpreting the findings.

Chapter 6 Discussion and conclusions

Summary and interpretation of main findings

The trial evaluated men following care pathways, extending from initial referral with bothersome LUTS to final symptom outcome. The two arms differed by randomisation to an additional test (UDS) in the diagnostic component of one arm. The trial established that the UDS arm is non-inferior to routine care for IPSSs at 18 months of follow-up, with a CI within the non-inferiority margin of 1 point. The IPSSs dropped to a similar extent in both arms [from an overall mean of 19 ($n = 774$) to a mean of 13 ($n = 669$)]. However, the anticipated potential advantage of lower surgery rates in the UDS arm was not found. Surgery rates, generally, in the trial were lower than expected and similar between the arms (38% and 36% for the UDS and routine care arms, respectively). Thus, our results do not support the routine use of UDS in men undergoing investigation of LUTS. However, it should not be extrapolated that UDS does not have a selective role to play for individual patients and for those patients with characteristics that led to exclusion from the trial.

Primary and key secondary outcomes

The mean baseline IPSS was 19 in both arms, with a good match of individual symptoms between the arms; 48% of men were categorised as having severe LUTS according to their IPSS. At 18 months, this had dropped to 21%. The trial confirmed that the arm in which UDS was included was non-inferior in terms of IPSS outcome at 18 months compared with routine care on an ITT basis, using the prespecified non-inferiority margin. In both arms, just over 80% of men completed their IPSS at the 18-month post-randomisation stage, as required in the initial sample size calculation.

It was hypothesised that the UDS arm would have a lower surgery rate, as the additional testing would be able to identify those men whose voiding LUTS were due to weakness of bladder contraction and thereby avoid surgery to reduce BOO. The actual finding was that the surgery rates did not differ between arms; this applied using ITT and actual treatment analysis. In practice, whether or not to proceed to surgery is decided in dialogue, based on the recommendation of the surgeon studying all the available information from the diagnostic tests, and the acceptance or otherwise by the man. On acceptance, the man is 'listed' for surgery; time between listing and admission to undergo surgery can influence the actual surgery rate (as opposed to the intention of surgery), as the progress was relatively slow in some centres. However, the number of people still waiting for surgery at trial completion was modest and did not affect the relative surgery rates. To focus on whether or not the surgeon recommendations in the trial may have reflected a surgeon response to the man's acceptance, an exercise was undertaken to derive treatment recommendations purely from UDS and free-flow rate/PVR parameters, using prespecified parameters agreed by the co-applicant urologists. This also found no difference between the arms, showing that the lack of difference between arms was a true reflection of test findings. Thus, the lack of difference in surgery rates is a clear-cut conclusion.

The interventions available for surgeons during the trial evolved over the time of the trial. The emergence of the UroLift procedure was clearly evident, as it was introduced during the course of the trial and was used to treat 7% of the men undergoing interventional treatment. Otherwise, the nature of the interventions used was reflective of the range of techniques employed in standard NHS practice when the trial was set up.

Secondary outcomes

The mean number of AEs experienced per person was similar between the arms. A possible preponderance of unrelated AEs was seen in the UDS arm, but this, as 'unrelated' suggests, is unlikely to be a consequence of diagnostic testing. Some AEs were specific to the diagnostic testing, but SAEs were not an issue for the tests, although little difference between the arms was evident. There were 11 deaths that were independent of trial procedures or interventions.

Patient-reported outcome measures for urinary symptoms showed a difference in the number of men reporting nocturia at 18 months, with fewer men reporting it in the UDS arm. PROMs for sexual symptoms showed no obvious differences between the arms at 18 months; however, there did appear to be a non-significant increase in the number of men suffering from reduced ejaculation in the UDS arm. Given the large number of secondary outcomes, and lack of evidence provided by other PROMs, these are unlikely to reflect a consequence of the diagnostic tests and could be chance findings.

A satisfaction questionnaire was specifically administered to men undergoing UDS, which showed high levels of satisfaction and hypothetical willingness to undergo repeat testing. Explanation of results was one area in which satisfaction was lower, compatible with the qualitative interview findings.[48]

Economic evaluation

The results of the economic analysis suggest that including UDS in the assessment of men who present with LUTS, compared with routine care, leads to higher costs and similar QALYs from all three perspectives (i.e. NHS secondary care, a wider NHS perspective and patient perspective). The wide and overlapping CIs around all the results show uncertainty. However, the results are robust to sensitivity analyses. CEACs demonstrate that the probability that UDS is cost-effective compared with routine care is 0.38 at a societal willingness to pay of £20,000 per QALY, rising to a probability of 0.6 if the willingness to pay is increased to £100,000 per QALY.

Qualitative evaluation

UPSTREAM achieved a large sample size, interviewing a range of trial participants in relation to demographics, presentation and pathway. To our knowledge, UPSTREAM is the first in-depth study of men's attitudes to and experiences of UDS for LUTS, using formal qualitative research methodology. It established the reasons why men were willing to participate in the trial, including conditional altruistic intent. It confirmed that men were willing to be randomised, although many men did express a preference for one of the arms. Those men unwilling to participate stated reasons, including disliking the intervention and the potential delay to treatment. As it turned out, UDS generally was well tolerated, with men experiencing minimal to moderate discomfort and embarrassment, particularly resulting from the intimacy of the procedure.

Little has been published about clinician views about UDS, with existing literature focusing on female urinary incontinence and the use of UDS before surgery. Clinicians seem to be polarised regarding the usefulness and current use of UDS. Part of the polarisation may reflect the extent to which clinicians are trained in the techniques and familiar with their interpretation. Another issue is lack of clarity in existing guidelines on how and when to use UDS.

Qualitative interviewing identified the need for clinicians to better inform men in advance about what to expect during and after UDS, including the risk of UTIs. In the routine care arm, 26 men received UDS, of whom 20 underwent UDS as a consequence of the clinician's decision to conduct UDS to aid their surgery decision. Qualitative interviews identified the attitude of clinicians in considering the importance of certain perceived indicators or risk in driving the requirement of UDS testing. There were more withdrawals

in the UDS arm than in the routine care arm, and some explanation for this difference might be a consequence of the perceived nature of UDS. Clinicians appeared to have some influence over the men's attitudes towards UDS, in the way that they describe the test itself and portray the potential benefit and harms.

Strength and limitations

A total of 8671 men were screened for eligibility, of whom 1482 (17%) were considered eligible and 820 (55%) were randomised; the prespecified sample size requirement (310 men per arm) was met with sufficient complete data sets. Adherence to allocation was reasonable; 83% of men allocated to the UDS arm underwent UDS. Of those men that did not undergo UDS, some declined to have the test and other men reported a resolution of their symptoms. Neither of these reasons was reported in the routine care arm. Satisfaction with UDS testing was high, although findings do not represent participants who declined to have the test (14 men), or those men who declined to take part in the trial because they did not want to be randomised (144/662 men) or did not want UDS or additional tests (41/662 men). There was a difference in the randomisation allocations, with 427 men in the UDS arm and 393 men in the routine care arm, which probably reflected the use of simple randomisation. This was discussed with the DMC and TSC during the course of the trial, and a specific check was made with the randomisation service to confirm that the discrepancy was compatible with the tolerance of the randomisation system and that it arose by chance. Our intention was to stratify randomisation by centre,[19] but an error meant that this was not undertaken (as reported in our published analysis plan[20]). Although this was a limitation, we did adjust for centre in all of our analyses, to account for some of the bias caused by hospital-specific treatment pathways.

Sixty-seven men withdrew from the trial before the 18-month end-of-trial visit (seven of whom requested complete data withdrawal). The reasons for withdrawal were broadly matched between the trial arms and reflect both real-life practice and typical research experience for a group of men in this age range, and with the extent of comorbidity. Comorbidities were reported in two-thirds of the men for both arms.

The trial arms were well matched in terms of demographic parameters. There were more withdrawals in the UDS arm, but only to a modest extent. None of the demographic parameters required a sensitivity analysis due to differences exceeding 10% (or a 0.5 SD). More than 90% of men in both arms of the trial were white and there was a broad range of deprivation scores, which were similar between the two arms. Flow rate testing results were similar for the two arms at the baseline. Q_{max} was 10.0 ml/second in the UDS arm and 10.9 ml/second in the routine care arm, which is entirely typical of a population of men with LUTS. PVRs were 95.0 ml and 90.0 ml, with VVs of 205 ml and 197 ml for the UDS and routine care arms, respectively. Again, these are compatible with the studied population. There was a considerable range of effectiveness in recruitment per centre.

The relatively low surgery rates are a clear feature and, in fact, there is an evident trend in reducing interventional therapies over recent years. At the start of the trial, all centres underwent a site set-up process. Such an experience clearly focusses the participating staff on the diagnostic pathway and provides an educational aspect. Although this was not the prime intention of the site set-up meeting, it may have improved the standard care in both arms simply by being involved in the trial. The more in-depth discussion between surgeon and patient may have provided an environment in which the clinician and patient could consider reasons not to list for surgery. There is recognised potential for behaviour change in both staff and patients in research studies. Indeed, our qualitative findings support the fundamental benefits of honest and detailed dialogue. Accordingly, some men managed conservatively and with minimal change in LUTS severity reported an improved QoL, with better knowledge and reassurance presumably contributing to some of the improvement. The crucial need for dialogue may be deduced from the acceptance rates; if the surgeon recommended surgery, fewer than 85% of men accepted it. In contrast, if the surgeon did not recommend surgery, more than 97% accepted it.

The strengths of the qualitative study include a focus on in-depth understanding of patients' and clinicians' perspectives and experiences of the trial, UDS and surgery. We were able to recruit a large and diverse sample of participants in terms of age, symptom burden and treatment decision. We were also able to speak to several health-care professionals in various centres. In both groups we successfully attained data saturation. We included participants with no direct experience of UDS, as the perceptions of these people can help establish acceptability in the target population. There was a high degree of similarity between pre-treatment and post-surgery interviews. Both identified positive and negative views and experiences of UDS testing, unrelated to the subsequent treatment and ultimate outcome. A weakness of the study is the possibility that the findings did not adequately cover the people who declined to participate in UPSTREAM, as they would never countenance undergoing UDS. Although we explored prior UDS practice at the participating sites in the clinician interviews, we did not interview any clinicians at baseline (prior to site initiation), when they might have provided different opinions about UDS and its practice.

As prespecified in the trial protocol and monitoring plan (summarised in *Appendix 7*), we reviewed equipment maintenance logs from 25 of the 26 centres and scrutinised trace interpretation and reporting to confirm compliance with ICS standards.[57,73] Although the majority of practice was broadly compliant with the standards, specific incidences were identified that were not. These were reviewed by the trial team to identify and remedy issues for effective delivery of the trial.

Overall evidence and generalisability

The full spectrum of potential LUTS was reported by men evaluated in UPSTREAM. Voiding symptoms were common, but storage symptoms and nocturia, in particular, were also highly prevalent. This reflects the real-life experience of urology departments, as the reported epidemiological studies point to the high levels of storage LUTS in the older male population.[74]

Recommendations for research

Each of the tests potentially used in the assessment of LUTS contributes elements that, taken together, facilitate an understanding of the mechanisms and hence the logical choice of therapy. There are a number of possible treatments that men may be offered based on these collective tests: watchful waiting, conservative treatment, pharmacological management or surgery. The most cost-effective diagnostic strategy in this type of diagnostic pathway could be assessed using a decision model.

In both arms, there was an overall reduction in IPSSs by 18 months, and having surgery led to the greatest reduction in IPSS. However, 147 men experienced a worsening of IPSS severity and 91 reported worsening of IPSS QoL. This may reflect natural progression of untreated LUTS, but some of the men underwent surgery. These men are a particular focus for further analysis, as the diagnostic precision may potentially be further optimised, to elicit factors that may anticipate complications or symptom deterioration. Identifying the predictive factors for worsening symptoms, or for strong symptomatic improvement, is the priority of the ongoing interpretation of findings to potentially target UDS to patient groups most likely to benefit from the procedure.

The long-term outcomes for men treated for LUTS remain unclear, particularly regarding storage LUTS such as nocturia. Beyond the 18-month time point, the need for ongoing or additional treatment is probable, particularly for those men who did not have surgery and for those men undergoing surgery whose underlying mechanism was DU.

The ICIQ-MLUTS scores showed changes in line with the IPSSs, which is unsurprising given that both are patient-reported outcomes for LUTS severity. However, there are differences between the two PROMs, which may provide a more detailed evaluation of symptoms, notably in the areas of urgency incontinence,

post-micturition dribble and bother measurements. Identifying how different PROMs work may establish whether or not they should be employed, preferentially for clinical use.

Implications for health care

The UPSTREAM findings do not support the routine use of UDS tests in men undergoing investigation of LUTS in the NHS; that is the principal finding of this trial. Further responder analysis of the UPSTREAM results will be used to show specific presentations in which individuals may require the additional information of UDS testing.

The expressed need of men is to seek relief of bothersome symptoms, as opposed to relief of BOO. Hence, the urology department evaluation needs to establish which LUTS are bothering individual men, to direct the focus of testing and treatment. This is particularly important for therapy selection, notably the use of surgery to relieve BOO; at presentation, the man needs to report bother from voiding LUTS to be able to expect improved QoL from a procedure whose principal aim is to improve voiding. If voiding LUTS are non-bothersome, then there should be an aspect that suggests that the bothersome LUTS could realistically improve following surgery. For example, a PVR contributing to bothersome frequency of passing urine or nocturia.

In theory, prostatic interventions aimed at reducing established BOO should directly improve voiding symptoms. Any benefit to storage symptoms would be indirect, perhaps driven by the reduced outlet resistance leading to an improved behaviour of the detrusor smooth muscle of the bladder following surgery. These indirect mechanisms of potential benefit need to be recognised as such, along with the implications of treatment AEs, which could add an extra bother component, requiring weighing in to the treatment response.

Any benefit of surgery to relieve BOO in terms of nocturia would be indirect as well, as nocturia is, in large part, determined by the production of urine during the night, which is a systemic factor for the handling of elimination of surplus water, salt and toxins. Both arms had a notably high prevalence of nocturia of more than once per night (75–80%) and impaired erectile function (71–76%). This real-life setting, in which common associated factors can influence decision-making and outcome, makes the UPSTREAM population very relevant to everyday practice.

This trial identified service delivery limitations, including equipment calibration and maintenance, interpretation of data and review of findings. Notably, we identified areas where the patients reported issues related to health-care contact, briefing, running tests, explanation of findings and shared decision-making.

Conclusions

Inclusion of UDS in the range of diagnostic tests for male LUTS results in a symptom outcome that is non-inferior to routine care. However, this does not have an impact on the rates of surgical procedures for treating BOO. Including UDS in the assessment pathway leads to higher costs and similar QALYs from the perspectives of NHS secondary care, the wider NHS and patients. The qualitative research identified that UDS was acceptable to patients, and valued by both patients and clinicians for its perceived additional insight into the cause and probable best treatment of LUTS. Overall, the routine use of UDS for all men with suspected BOO is not supported by the study findings. However, the large number of men who saw modest symptom change, or evident deterioration, does not exclude selective use of UDS for specific individuals, which we propose exploring further.

Acknowledgements

The authors wish to thank all the men who participated in UPSTREAM. We would also like to thank all the participating NHS hospitals (centres), 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[19]). The trial 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] and the DMC [Mr Matt Sydes, Reader in Clinical Trials (chairperson); Miss Suzanne Biers, Consultant Urological Surgeon; and Mr Nikesh Thiruchelevam, Consultant Urological Surgeon], as well as the TMG (consisting of co-applicants, representatives from the trial office and patient representatives) for their continued counsel and support.

We are also grateful to Lyndsey Johnson, Samantha Clarke and Kwame Ansu for their contributions as UPSTREAM lead research nurses at different periods throughout the trial, and Charlotte McDonald and Aneta Taylor for their administrative support in recent years.

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 trial design and oversight.

Finally, we thank the NIHR and the HTA programme for funding UPSTREAM (reference number 12/140/01), and NBT for trial sponsorship (reference number 3250).

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

Trial data were collected and managed using REDCap,[45] hosted at the University of Bristol.

Contributions of authors

Amanda L Lewis (Trial Manager) was responsible for managing the trial, contributed to aspects of trial design and the SAP, and prepared and contributed to writing the report.

Grace J Young (Senior Research Associate, Medical Statistics) was the trial statistician involved with writing the SAP and conducting the analysis, and contributed to writing the report (in particular *Chapters 2* and *3*).

Lucy E Selman (Senior Research Fellow in Qualitative Research in Randomised Trials) contributed to the design, management and conduct of the qualitative study, and contributed to writing the report (in particular *Chapter 5*).

Caoimhe Rice (Research Associate, Health Economics) was involved in the design, conduct and analysis of the economic evaluation, and contributed to writing the report (in particular *Chapter 4*).

Clare Clement (Senior Research Associate in Qualitative Research in Randomised Trials) was involved in the conduct of the qualitative study and contributed to writing the report (in particular *Chapter 5*).

Cynthia A Ochieng (Postdoctoral Qualitative Researcher) was involved in the conduct of the qualitative study and contributed to writing the report (in particular *Chapter 5*).

Publications

Journal articles

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.

Lewis AL, Drake MJ. Urodynamics for Prostate Surgery Trial; Randomised Evaluation of Assessment Methods (UPSTREAM) for diagnosis and management of bladder outlet obstruction in men. *J Clin Urol* 2015;**8**:296–8.

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.

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.

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.

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.

Aiello M, Jelksi J, Lewis AL, 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.

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 [published online ahead of print June 30 2020]. *Eur Urol* 2020.

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; in press.

Conference abstract

Rice C, Lewis AL, Noble S, Drake MJ. *Direct Data Without Duplication of Effort: Experience of Obtaining Electronic Routine Resource Use Data Directly From Hospitals From the UPSTREAM RCT.* International Society for Pharmacoeconomics and Outcomes Research, Inc. (ISPOR) Europe, Barcelona, Spain, 13 November 2018, abstract no. 96.

Drake M, Abrams PH, Chapple C, Lewis A. The diagnostic assessment pathway for male LUTS; a randomised trial of urodynamics or usual care in 26 hospitals. Rome: 17th Meeting of the Association of Academic European Urologists; 30 November 2018.

Drake M. What is the optimal treatment for patients with male LUTS? Barcelona: European Association of Urology Annual Meeting (EAU); 19 March 2019.

Drake M. The UPSTREAM study. Glasgow: The British Association of Urological Surgeons (BAUS) Annual Meeting; 26 June 2019.

Horwood J, Clement C, Ochieng C, Lewis AL, Chapple C, Abrams P, *et al.* The value of qualitative methods in interpreting trial outcomes to inform clinical practice: understanding UPSTREAM trial participants treatment decision making for lower urinary tract symptoms. Brighton: 5th International Clinical Trials Methodology Conference (ICTMC); 6–9 October 2019, abstract no. P-123.

Rice C, Lewis AL, Noble S, Drake MJ. Obtaining electronic routine resource use data directly from hospitals; experience from the UPSTREAM randomised controlled trial. Brighton: 5th International Clinical Trials Methodology Conference (ICTMC); 6–9 October 2019, abstract no. P-47.

Data-sharing statement

Reasonable data-sharing requests should be made to the corresponding author. Requests will be reviewed with advisement of the TMG and trial sponsor (NBT).

Patient data

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's important that there are safeguards to make sure that it is 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.

References

1. 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. https://doi.org/10.1016/S0022-5347(01)66461-9

2. Orlowski A, Kayes O. Comparison of costs associated with TURP and prostatic urethral lift for benign prostatic hyperplasia. *Value Health* 2018;**21**:S268. https://doi.org/10.1016/j.jval.2018.04.1811

3. Rassweiler J, Teber D, Kuntz R, Hofmann R. Complications of transurethral resection of the prostate (TURP) – incidence, management, and prevention. *Eur Urol* 2006;**50**:969–79. https://doi.org/10.1016/j.eururo.2005.12.042

4. Gratzke C, Bachmann A, Descazeaud A, Drake MJ, Madersbacher S, Mamoulakis C, *et al.* EAU guidelines on the assessment of non-neurogenic male lower urinary tract symptoms including benign prostatic obstruction. *Eur Urol* 2015;**67**:1099–109. https://doi.org/10.1016/j.eururo.2014.12.038

5. National Institute for Health and Care Excellence (NICE). *Lower Urinary Tract Symptoms in Men: Management [CG97]*. London: NICE; 2010. URL: www.nice.org.uk/guidance/cg97 (accessed 25 July 2019).

6. Reynard JM, Peters TJ, Lim C, Abrams P. The value of multiple free-flow studies in men with lower urinary tract symptoms. *Br J Urol* 1996;**77**:813–18. https://doi.org/10.1046/j.1464-410X.1996.00097.x

7. Foster HE, Barry MJ, Dahm P, Gandhi MC, Kaplan SA, Kohler TS, *et al.* Surgical management of lower urinary tract symptoms attributed to benign prostatic hyperplasia: AUA guideline. *J Urol* 2018;**200**:612–19. https://doi.org/10.1016/j.juro.2018.05.048

8. Abrams PH, Griffiths DJ. The assessment of prostatic obstruction from urodynamic measurements and from residual urine. *Br J Urol* 1979;**51**:129–34. https://doi.org/10.1111/j.1464-410X.1979.tb02846.x

9. Abrams P. Bladder outlet obstruction index, bladder contractility index and bladder voiding efficiency: three simple indices to define bladder voiding function. *BJU Int* 1999;**84**:14–15. https://doi.org/10.1046/j.1464-410x.1999.00121.x

10. Clement KD, Burden H, Warren K, Lapitan MC, Omar MI, Drake MJ. Invasive urodynamic studies for the management of lower urinary tract symptoms (LUTS) in men with voiding dysfunction. *Cochrane Database Syst Rev* 2015;**4**:CD011179. https://doi.org/10.1002/14651858.CD011179.pub2

11. Parsons BA, Bright E, Shaban AM, Whitehouse A, Drake MJ. The role of invasive and non-invasive urodynamics in male voiding lower urinary tract symptoms. *World J Urol* 2011;**29**:191–7. https://doi.org/10.1007/s00345-009-0488-8

12. Chapple CR, Osman NI, Birder L, Dmochowski R, Drake MJ, van Koeveringe G, *et al.* Terminology report from the International Continence Society (ICS) Working Group on underactive bladder (UAB). *Neurourol Urodyn* 2018;**37**:2928–31. https://doi.org/10.1002/nau.23701

13. Drake MJ. Do we need a new definition of the overactive bladder syndrome? ICI-RS 2013. *Neurourol Urodyn* 2014;**33**:622–4. https://doi.org/10.1002/nau.22609

14. Abrams P, Cardozo L, Fall M, Griffiths D, Rosier P, Ulmsten U, *et al.* The standardisation of terminology of lower urinary tract function: report from the Standardisation Sub-committee of the International Continence Society. *Neurourol Urodyn* 2002;**21**:167–78. https://doi.org/10.1002/nau.10052

15. Van Kerrebroeck P, Abrams P, Chaikin D, Donovan J, Fonda D, Jackson S, *et al.* The standardization of terminology in nocturia: report from the standardization subcommittee of the International Continence Society. *BJU Int* 2002;**90**(Suppl. 3):11–15. https://doi.org/10.1046/j.1464-410X.90.s3.3.x

16. Drake MJ. Should nocturia not be called a lower urinary tract symptom? *Eur Urol* 2015;**67**:289–90. https://doi.org/10.1016/j.eururo.2014.09.024

17. Bosch JL, Everaert K, Weiss JP, Hashim H, Rahnama'i MS, Goessaert AS, Aizen J. Would a new definition and classification of nocturia and nocturnal polyuria improve our management of patients? ICI-RS 2014. *Neurourol Urodyn* 2016;**35**:283–7. https://doi.org/10.1002/nau.22772

18. Reynard JM, Yang Q, Donovan JL, Peters TJ, Schafer W, de la Rosette JJ, *et al.* The ICS-'BPH' Study: uroflowmetry, lower urinary tract symptoms and bladder outlet obstruction. *Br J Urol* 1998;**82**:619–23. https://doi.org/10.1046/j.1464-410X.1998.00813.x

19. 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

20. 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

21. Department of Health and Social Care. *Research Governance Framework for Health and Social Care*. London: Department of Health and Social Care; 2005. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/139565/dh_4122427.pdf (accessed 16 October 2018).

22. World Medical Association. *WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*. Ferney-Voltaire: The World Medical Association, Inc.; 2018. URL: www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ (accessed 16 October 2018).

23. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 2004;**240**:205–13. https://doi.org/10.1097/01.sla.0000133083.54934.ae

24. Barry MJ, Cantor A, Roehrborn CG, CAMUS Study Group. Relationships among participant international prostate symptom score, benign prostatic hyperplasia impact index changes and global ratings of change in a trial of phytotherapy in men with lower urinary tract symptoms. *J Urol* 2013;**189**:987–92. https://doi.org/10.1016/j.juro.2012.08.257

25. Tikkinen KA, Johnson TM, Tammela TL, Sintonen H, Haukka J, Huhtala H, Auvinen A. Nocturia frequency, bother, and quality of life: how often is too often? A population-based study in Finland. *Eur Urol* 2010;**57**:488–96. https://doi.org/10.1016/j.eururo.2009.03.080

26. Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. *Trials* 2011;**12**:106. https://doi.org/10.1186/1745-6215-12-106

27. 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. https://doi.org/10.1016/j.euf.2019.04.006

28. 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 [published online ahead of print June 30 2020]. *Eur Urol* 2020. https://doi.org/10.1016/j.eururo.2020.06.004

29. Aiello M, Jelksi J, Lewis AL, 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

30. Manchester Information and Associated Services (MIMAS). *GeoConvert*. URL: http://geoconvert.mimas.ac.uk/ (accessed 19 October 2018).

31. 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. https://doi.org/10.1016/j.jval.2012.02.008

32. NHS Digital. *NHS Reference Costs Grouper*. Leeds: NHS Digital; 2017. URL: https://digital.nhs.uk/services/national-casemix-office/downloads-groupers-and-tools/grouper-and-tools-archive/costing-hrg4-2016-17-reference-costs-grouper (accessed 30 July 2018).

33. NHS Digital. *NHS Reference Costs 2016/17*. Leeds: NHS Digital; 2017. URL: https://improvement.nhs.uk/resources/reference-costs/ (accessed 8 August 2018).

34. Curtis L, Burns A. *Unit Costs of Health and Social Care 2017*. Canterbury: Personal Social Services Research Unit, University of Kent; 2017. https://doi.org/10.22024/UniKent/01.02/65559

35. Curtis L. *Unit Costs of Health and Social Care 2013*. Canterbury: Personal Social Services Research Unit, University of Kent; 2013.

36. Curtis L, Burns A. *Unit Costs of Health and Social Care 2015*. Canterbury: Personal Social Services Research Unit, University of Kent; 2015.

37. GOV.UK. *Expenses and Benefits: Business Travel Mileage for Employees' Own Vehicles*. URL: www.gov.uk/expenses-and-benefits-business-travel-mileage/rules-for-tax (accessed 25 October 2018).

38. NHS Commissioning Board (North of England Commissioning Support). *Increase to NHS Prescription Charges From 1st April 2016*. URL: https://medicines.necsu.nhs.uk/increase-to-nhs-prescription-charges-from-1st-april-2016/ (accessed 25 October 2018).

39. Office for National Statistics. *National Minimum Wage and National Living Wage Rates*. Newport: Office for National Statistics. URL: www.gov.uk/national-minimum-wage-rates (accessed 24 September 2018).

40. National Audit Office. *Stocktake of Access to General Practice in England*. London: Department of Health and NHS England; 2015. URL: www.nao.org.uk/wp-content/uploads/2015/11/Stocktake-of-access-to-general-practice-in-England.pdf (accessed 18 October 2018).

41. Roberts A, Blunt I, Ariti C, Bardsley M. *Focus on: Distance From Home for Emergency Care. Research Report*. London: Nuffield Trust and Health Foundation; 2014.

42. National Institute for Health and Care Excellence. *The Guidelines Manual*. London: National Institute for Health and Care Excellence; 2012. URL: www.nice.org.uk/process/pmg6/chapter/assessing-cost-effectiveness (accessed 20 September 2018).

43. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;**30**:377–99. https://doi.org/10.1002/sim.4067

44. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;**14**:487–96. https://doi.org/10.1002/hec.944

45. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;**42**:377–81. https://doi.org/10.1016/j.jbi.2008.08.010

46. Marques E, Johnson EC, Gooberman-Hill R, Blom AW, Noble S. Using resource use logs to reduce the amount of missing data in economic evaluations alongside trials. *Value Health* 2013;**16**:195–201. https://doi.org/10.1016/j.jval.2012.09.008

47. Heslin M, Babalola O, Ibrahim F, Stringer D, Scott D, Patel A. A comparison of different approaches for costing medication use in an economic evaluation. *Value Health* 2018;**21**:185–92. https://doi.org/10.1016/j.jval.2017.02.001

48. 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. https://doi.org/10.1002/nau.23855

49. Medical Research Council. A *Framework for Development and Evaluation of RCTs for Complex Interventions to Improve Health*. Swindon: Medical Research Council; 2000. URL: www.mrc.ac.uk/documents/pdf/rcts-for-complex-interventions-to-improve-health/ (accessed 17 August 2015).

50. Donovan J, Mills N, Smith M, Brindle L, Jacoby A, Peters T, *et al.* Quality improvement report: improving design and conduct of randomised trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study. Commentary: presenting unbiased information to patients can be difficult. *BMJ* 2002;**325**:766–70. https://doi.org/10.1136/bmj.325.7367.766

51. Barry MJ, Fowler FJ, O'leary MP, Bruskewitz RC, Holtgrewe HL, Mebust WK, Cockett AT, Measurement Committee of the American Urological Association. The American Urological Association symptom index for benign prostatic hyperplasia. *J Urol* 2017;**197**:S189–S197. https://doi.org/10.1016/j.juro.2016.10.071

52. Smith TNM, Noble S, Wright G, McLennan D, Plunkett E. *The English Indices of Deprivation 2015*. London: Department for Communities and Local Government; 2015. URL: www.gov.uk/government/statistics/english-indices-of-deprivation-2015 (accessed 3 October 2018).

53. Sandelowski M. Sample size in qualitative research. *Res Nurs Health* 1995;**18**:179–83. https://doi.org/10.1002/nur.4770180211

54. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. https://doi.org/10.1191/1478088706qp063oa

55. Mills N, Donovan JL, Wade J, Hamdy FC, Neal DE, Lane JA. Exploring treatment preferences facilitated recruitment to randomized controlled trials. *J Clin Epidemiol* 2011;**64**:1127–36. https://doi.org/10.1016/j.jclinepi.2010.12.017

56. Ellis-Jones J, Swithinbank L, Abrams P. The bridges and barriers to 'good' urodynamic practice: a regional perspective. *Int J Urol Nurs* 2013;**7**:3–8. https://doi.org/10.1111/j.1749-771X.2012.01154.x

57. Rosier PFWM, Schaefer W, Lose G, Goldman HB, Guralnick M, Eustice S, *et al.* International Continence Society Good Urodynamic Practices and Terms 2016: urodynamics, uroflowmetry, cystometry, and pressure-flow study. *Neurourol Urodyn* 2017;**36**:1243–60. https://doi.org/10.1002/nau.23124

58. Glover L, Gannon K, McLoughlin J, Emberton M. Men's experiences of having lower urinary tract symptoms: factors relating to bother. *BJU Int* 2004;**94**:563–7. https://doi.org/10.1111/j.1464-410X.2004.05001.x

59. Gannon K, Glover L, O'Neill M, Emberton M. Men and chronic illness: a qualitative study of LUTS. *J Health Psychol* 2004;**9**:411–20. https://doi.org/10.1177/1359105304042350

60. Wareing M. Lower urinary tract symptoms: a hermeneutic phenomenological study into men's lived experience. *J Clin Nurs* 2005;**14**:239–46. https://doi.org/10.1111/j.1365-2702.2004.01003.x

61. Welch LC, Botelho EM, Tennstedt SL. Race and ethnic differences in health beliefs about lower urinary tract symptoms. *Nurs Res* 2011;**60**:165–72. https://doi.org/10.1097/NNR.0b013e3182159cac

62. Shaw C, Tansey R, Jackson C, Hyde C, Allan R. Barriers to help seeking in people with urinary symptoms. *Fam Pract* 2001;**18**:48–52. https://doi.org/10.1093/fampra/18.1.48

63. Coyne KS, Sexton CC, Kopp Z, Chapple CR, Kaplan SA, Aiyer LP, Symonds T. Assessing patients' descriptions of lower urinary tract symptoms (LUTS) and perspectives on treatment outcomes: results of qualitative research. *Int J Clin Pract* 2010;**64**:1260–78. https://doi.org/10.1111/j.1742-1241.2010.02450.x

64. Oh SJ, Son H, Jeong JY, Ku JH. Patients' experience with ambulatory urodynamics. A prospective study. *Scand J Urol Nephrol* 2006;**40**:391–6. https://doi.org/10.1080/00365590600744014

65. Almallah YZ, Rennie CD, Stone J, Lancashire MJ. Urinary tract infection and patient satisfaction after flexible cystoscopy and urodynamic evaluation. *Urology* 2000;**56**:37–9. https://doi.org/10.1016/S0090-4295(00)00555-0

66. Suskind AM, Clemens JQ, Kaufman SR, Stoffel JT, Oldendorf A, Malaeb BS, *et al.* Patient perceptions of physical and emotional discomfort related to urodynamic testing: a questionnaire-based study in men and women with and without neurologic conditions. *Urology* 2015;**85**:547–51. https://doi.org/10.1016/j.urology.2014.11.001

67. Yokoyama T, Nozaki K, Nose H, Inoue M, Nishiyama Y, Kumon H. Tolerability and morbidity of urodynamic testing: a questionnaire-based study. *Urology* 2005;**66**:74–6. https://doi.org/10.1016/j.urology.2005.01.027

68. Quek P, Tay LH. Morbidity and significant bacteriuria after urodynamic studies. *Ann Acad Med Singap* 2004;**33**:754–7.

69. Scarpero HM, Padmanabhan P, Xue X, Nitti VW. Patient perception of videourodynamic testing: a questionnaire based study. *J Urol* 2005;**173**:555–9. https://doi.org/10.1097/01.ju.0000149968.60938.c0

70. Shaw C, Williams K, Assassa PR, Jackson C. Patient satisfaction with urodynamics: a qualitative study. *J Adv Nurs* 2000;**32**:1356–63. https://doi.org/10.1046/j.1365-2648.2000.01627.x

71. Yiou R, Audureau E, Loche CM, Dussaud M, Lingombet O, Binhas M. Comprehensive evaluation of embarrassment and pain associated with invasive urodynamics. *Neurourol Urodyn* 2015;**34**:156–60. https://doi.org/10.1002/nau.22521

72. Hilton P, Bryant A, Howel D, McColl E, Buckley BS, Lucas M, *et al.* Assessing professional equipoise and views about a future clinical trial of invasive urodynamics prior to surgery for stress urinary incontinence in women: a survey within a mixed methods feasibility study. *Neurourol Urodyn* 2012;**31**:1223–30. https://doi.org/10.1002/nau.22328

73. Gammie A, Clarkson B, Constantinou C, Damaser M, Drinnan M, Geleijnse G, *et al.* International Continence Society guidelines on urodynamic equipment performance. *Neurourol Urodyn* 2014;**33**:370–9. https://doi.org/10.1002/nau.22546

74. Coyne KS, Sexton CC, Thompson CL, Milsom I, Irwin D, Kopp ZS, *et al.* The prevalence of lower urinary tract symptoms (LUTS) in the USA, the UK and Sweden: results from the Epidemiology of LUTS (EpiLUTS) study. *BJU Int* 2009;**104**:352–60. https://doi.org/10.1111/j.1464-410X.2009.08427.x

75. Schäfer W, Abrams P, Liao L, Mattiasson A, Pesce F, Spangberg A, *et al.* Good urodynamic practices: uroflowmetry, filling cystometry, and pressure-flow studies. *Neurourol Urodyn* 2002;**21**:261–74. https://doi.org/10.1002/nau.10066

Appendix 1 Overview of patient screening

TABLE 34 Summary of screening logs

Screening outcome	Number of patients ($n = 8671^a$)	%
Of those patients screened		
Deemed ineligible	5910	68
Deemed eligible	1482	17
Reasons for non-inclusion unidentified	1279	15
Of those patients confirmed as eligible		
Declined to take part	662	45
Randomised	820	55

a Total patients screened across 26 centres.
Reprinted from *European Urology Focus*, vol. 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). pp. 340–50, copyright 2019,[27] with permission from Elsevier.

TABLE 35 Summary of reasons why men who were eligible declined to take part ($n = 662$)

Reasons	Number of patients ($n = 662$)	%
Reasons	535	81
Does not want to be randomised	144	27
Other commitments	89	17
Could not decide	79	15
Number of visits	76	14
Not interested in research studies	56	10
Does not want UDS or additional tests	41	8
Other health issues more important	29	5
Number of questionnaires	16	3
Transport/parking issues	3	1
Relocating	2	0
Reason missing	127	19

Reprinted from *European Urology Focus*, vol. 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). pp. 340–50, copyright 2019,[27] with permission from Elsevier.

TABLE 36 Summary of reasons why men were ineligible to take part ($n = 5910$)

Reasons	Number of patients ($n = 5910$)	%
Exclusion criteria	2926	50
Undergoing treatment/surveillance prostate or bladder cancer	1293	44
Previous prostate surgery	559	19
Urinary retention	392	13
Neurological disease	302	10
Not willing/able to comply with essential study procedures	208	7
Not medically fit for surgery	172	6
Other reasons	2510	42
Medical team did not consider patient suitable for research/this study (non-prostatic/non-bothersome LUTS; presentation required additional assessment; recurrent UTIs; unrelated condition)	1068	43
Further details not provided	714	28
Patient no longer seeking treatment (or surgery)	346	14
Already had diagnostic assessments (e.g. UDS) and/or treatment plan is active (e.g. surgery/medication)	341	14
Considered too young, especially for surgery	35	1
Unable to make further contact with patient: deemed ineligible	6	0
Reason missing	474	8

Reprinted from *European Urology Focus*, vol. 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). pp. 340–50, copyright 2019,[27] with permission from Elsevier.

Appendix 2 Subgroup interactions with effectiveness of surgery

The effect of surgery on IPSSs seemed to be mediated by age. When looking at the effect modification of age on the relationship between surgery and IPSS, we see that surgery appears less effective in the older age groups and more effective in the younger age groups (*Figure 13*).

There did not appear to be any other effect modifiers when looking at other baseline variables mentioned in the EAU guidelines[4] (*Table 37*).
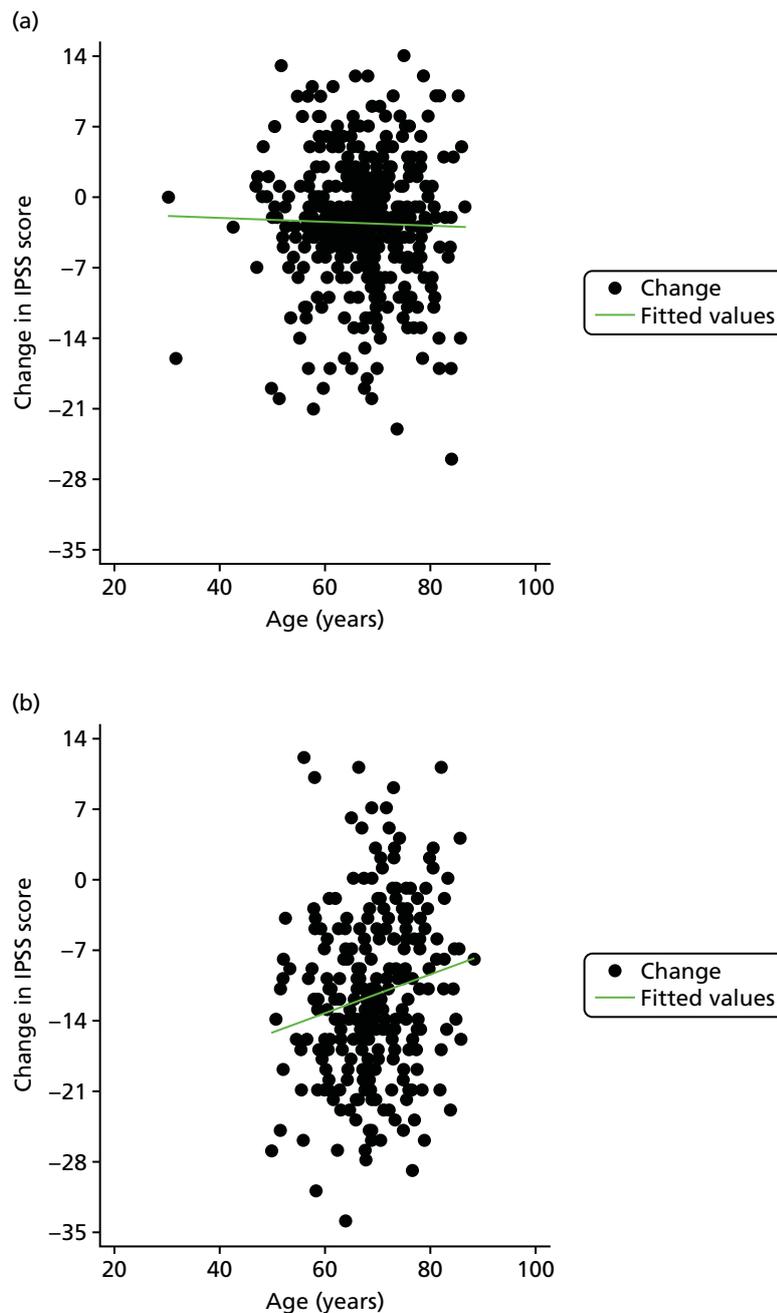

FIGURE 13 Subgroup interaction with effectiveness of surgery in (a) those that did not receive surgery; and (b) those that did receive surgery.

TABLE 37 Interaction between surgery and baseline cut-off points

Variable	n (surgery : no surgery)[a]	IPSS at 18 months[b] Subgroup-specific difference in means (95% CI)	Interaction effect[b] Difference in means (95% CI)	p-value
Subgroup analyses *Age (years)[c]*				
< 60	31 : 86	–8.94 (–12.07 to –5.81)	Baseline	
60–80	202 : 277	–7.25 (–8.53 to –5.97)	1.89 (–1.04 to 4.83)	0.205
> 80	20 : 25	–1.60 (–8.52 to 5.33)	6.34 (1.72 to 10.96)	0.007
VV (ml)[d]				
< 150	73 : 90	–6.92 (–9.33 to –4.51)	Baseline	
≥ 150	177 : 286	–7.44 (–8.78 to –6.09)	–0.60 (–2.98 to 1.78)	0.621
PVR (ml)[d]				
≤ 300	217 : 348	–7.06 (–8.27 to –5.85)	Baseline	
> 300	27 : 30	–10.48 (–14.66 to –6.30)	–1.44 (–4.98 to 2.09)	0.423
Nocturia				
Once (or not at all) per night	39 : 103	–7.85 (–10.18 to –5.51)	Baseline	
More than once per night	208 : 276	–7.43 (–8.76 to –6.10)	–0.68 (–3.33 to 1.96)	0.611

a Number per subgroup among those men who had surgery and those who did not have surgery.
b Adjusted for centre and baseline IPSS.
c Treating age as a continuous variable gave an interaction *p*-value of 0.002.
d Treating these variables as continuous did not affect the interaction *p*-value.

Appendix 3 Reasons for surgery recommendation

*T*able 38 shows the reasons given by the surgeons for either recommending or not recommending surgery. As this was based on ITT groups, there are some cases for which BOO was confirmed in the routine care arm, but this was largely because there was crossover between the arms. The most common reason for recommending surgery in the UDS arm was the confirmation of BOO being present (51%). However, for the routine care arm, the most common reason was unimproved symptoms (42%). A common reason in the UDS arm was no evidence of BOO (26%), which was most probably established from the UDS procedure. Other reasons were relatively balanced between the arms.

When categorising the reasons provided, the team were concerned that some of the recommendations may have been inconsistent with assessment findings and very variable across centres. To assess the benefits of UDS in the pathway, a post hoc analysis looked through each individual's baseline flow rate, VV, PVR overall IPSS, IPSS QoL, IPSS urgency question (question 4) and IPSS weak stream question (question 5). When men had UDS, the uroflowmetry results were scrutinised alongside the BCI and BOOI. When available, UDS results were used to make an initial treatment decision; when missing, flow rates were used. Criteria for recommending surgery consisted of the following, based on a consultation exercise of urologists not participating in UPSTREAM.

Surgery was recommended if:

- BOOI score of > 40; or Q_{max} of < 10 ml/second; or Q_{max} of between 10 and 15 ml/second; and PVR of > 100 ml.

Surgery was not recommended if:

- BOOI score of < 40 and BCI score of < 100; or BOOI score of < 20; or Q_{max} of > 15 ml/second.

Using these criteria, the recommendations would have been those displayed in *Table 39.*

Looking at these figures on an 'as-treated' basis gave very similar results.

TABLE 38 Considerations of surgical recommendations

	Primary reason		Secondary reason	
Surgeon recommendation	**UDS, *n* (%)**	**Routine care, *n* (%)**	**UDS, *n* (%)**	**Routine care, *n* (%)**
Surgeon recommended surgery				
BOO confirmed	99 (51)	7 (4)	0 (0)	0 (0)
Symptoms bothersome and unimproved	29 (15)	77 (42)	0 (0)	2 (6)
Storage LUTS	2 (1)	2 (1)	2 (6)	1 (3)
Appearance on endoscopic evaluation	6 (3)	3 (2)	0 (0)	0 (0)
Slow flow	16 (8)	24 (13)	0 (0)	5 (15)
PVR	6 (3)	7 (4)	19 (61)	17 (50)
IDUC	1 (1)	3 (2)	0 (0)	1 (3)
BPE	2 (1)	7 (4)	1 (3)	2 (6)
DU	1 (1)	0 (0)	1 (3)	0 (0)
Medical/conservative management not fully trialled	1 (1)	0 (0)	0 (0)	0 (0)

TABLE 38 Considerations of surgical recommendations (*continued*)

Surgeon recommendation	Primary reason		Secondary reason	
	UDS, *n* (%)	Routine care, *n* (%)	UDS, *n* (%)	Routine care, *n* (%)
Prevent retention	0 (0)	1 (1)	0 (0)	1 (3)
UTIs	1 (1)	0 (0)	0 (0)	0 (0)
Reassessed: surgery not indicated	2 (1)	0 (0)	0 (0)	0 (0)
Patient wants to stop medication	0 (0)	1 (1)	0 (0)	3 (9)
DO	0 (0)	0 (0)	8 (26)	1 (3)
DOI	0 (0)	0 (0)	0 (0)	1 (3)
Unclear reasoning	13 (7)	17 (9)	0 (0)	0 (0)
No reason stated	17 (9)	33 (18)	0 (0)	0 (0)
Surgeon did not recommend surgery				
No evidence of BOO	52 (26)	0 (0)	0 (0)	0 (0)
BOO confirmed	0 (0)	0 (0)	2 (11)	0 (0)
Symptoms bothersome and unimproved	0 (0)	0 (0)	0 (0)	0 (0)
Symptoms improved and not bothersome	67 (33)	86 (44)	0 (0)	1 (8)
Medical/conservative management not fully trialled	11 (5)	25 (13)	0 (0)	1 (8)
Patient does not want surgery	5 (2)	4 (2)	1 (6)	0 (0)
Storage LUTS	6 (3)	21 (11)	3 (17)	7 (54)
DU	9 (4)	0 (0)	4 (22)	0 (0)
Appearance on endoscopic evaluation	1 (< 1)	0 (0)	0 (0)	0 (0)
Good flow rate and/or low PVR	5 (2)	19 (10)	0 (0)	3 (23)
Screen for PCa	1 (< 1)	0 (0)	0 (0)	0 (0)
Unclear reasoning	6 (3)	22 (11)	0 (0)	0 (0)
No reason stated	19 (10)	17 (9)	0 (0)	0 (0)
DOI	1 (< 1)	0 (0)	0 (0)	0 (0)
Comorbidities	2 (1)	2 (1)	1 (6)	1 (8)
DO	13 (6)	0 (0)	7 (39)	0 (0)

DO, detrusor overactivity; DOI, detrusor overactivity incontinence; IDUC, indwelling urinary catheter; PCa, prostate cancer.

TABLE 39 Applying an internal universal surgery recommendation based on the available data

Randomisation group	UDS, *n* (%)	Routine care, *n* (%)
Had no UDS results been available[a]		
Recommended surgery	233 (70)	206 (66)
Not recommended surgery	100 (30)	104 (34)
Surgery recommendation based on the internal review[b]		
Recommended surgery	249 (65)	209 (66)
Not recommended surgery	136 (35)	106 (34)

a Using only flow rate and patient-reported data, such as Q_{max}, VV, PVR and IPSS questionnaires.
b Using UDS results in the first instance and then utilising flow rate and patient-reported data when UDS data were unavailable or unhelpful.
Note
Patients for whom the conclusion was 'maybe' or 'non-diagnostic' were not included in this table.

Appendix 4 Post-surgery urinary flow rate

Men had a surgery decision appointment in both arms of the trial. The clinician either recommended surgery or did not, based on the diagnostic test results available to them. For those men who went on to have surgery, a post-surgery flow rate test was carried out approximately 4 months after surgery ($\pm$ 1 month), with results contained in a postoperative CRF. Baseline and post-surgery flow rates are detailed in *Table 40*.

Descriptively, when looking at those men who were not recommended for surgery but ended up having it, results were not favourable. Again, the small sample size here makes it difficult to form any strong conclusions. When analysed by arm, there was no evidence to suggest that there was a difference in post-surgery Q_{max} (*Table 41*).

TABLE 40 Maximum urinary flow rates by treatment type

Group	UDS			Routine care		
	n (%)	Baseline, mean (SD)	Post surgery, mean (SD)	*n* (%)	Baseline, mean (SD)	Post surgery, mean (SD)
Had surgery for LUTS						
Had surgery for LUTS (all)	59	11.21 (7.12)	17.23 (9.28)	58	11.08 (5.12)	18.26 (9.86)
Recommended for surgery[a]	57	11.03 (7.11)	17.31 (9.42)	56	11.01 (5.20)	18.48 (9.96)
Not recommended for surgery[a]	2	16.35 (7.42)	15.05 (4.31)	2	13.15 (1.63)	12.15 (4.03)
TURP (any type)	44	10.77 (6.20)	18.39 (9.80)	49	11.33 (4.92)	18.52 (10.22)
Laser (any type)	8	15.24 (12.34)	17.70 (6.89)	4	8.90 (5.78)	18.50 (9.93)
BNI (with or without TURP)	1	6.70 (0.00)	3.9 (0.00)	2	5.30 (1.84)	19.00 (8.49)
UroLift	5	9.64 (3.65)	10.02 (1.66)	2	16.30 (8.91)	16.30 (4.67)

a Surgeon recommendation, decided during the decision-making appointment.

TABLE 41 Post-surgery flow rate analysis

Variable	*n* (UDS : routine care)	UDS, mean (SD)	Routine care, mean (SD)	Difference in means[a] (95% CI)	*p*-value[a]
Q_{max}					
Post surgery	59 : 58	17.23 (9.28)	18.26 (9.86)	–0.17 (–4.24 to 3.89)	0.932

a Adjustment for centre and baseline urinary flow rate.

Appendix 5 Key topic areas explored in trial participant interviews

All trial participants

Background

- Age, family situation, accommodation and work circumstances, general health.

Lower urinary tract symptoms experience

- Help-seeking triggers for LUTS.
- Symptoms experienced, bother and impact, primary concerns.
- Understanding and management of symptoms.
- Information received and needed in relation to LUTS.

Lower urinary tract symptoms assessment/consultation/treatments

- Patient pathway to hospital, referral process.
- Experiences at the hospital.
- Assessments and tests received for LUTS, including UDS testing, experiences, understanding of the assessment's purpose, expectations of assessments and information/support.
- Explanation of assessment results: format, adequacy, understanding the results.

Decision-making

- Treatment decision-making process and outcome.
- Patient and clinician involvement and role in decision-making.
- Patient preferences regarding treatment.
- Impact of assessments on treatment decision-making.
- Views of surgery for LUTS.

UPSTREAM

- Motivations for taking part in UPSTREAM, experiences of trial participation.
- Understanding of randomisation and equipoise, preferences regarding trial arm allocation.

Post-surgery trial participants

Experience of surgery, symptoms and recovery

- Expectations of surgery.
- Experience of surgery.
- Experience of recovery.
- Symptoms post surgery.

Appendix 6 Key topic areas explored in clinician interviews

Background

- Age, job title, year qualified, length of time working in urology, role.
- Details of routine care in relation to UDS.

UPSTREAM experience, context and pathway

- Involvement in UPSTREAM.
- Views and experiences of the trial.

Views of urodynamic testing

- Views before the trial and after, acceptability.
- Experience of conducting UDS.
- Impact of UDS on the treatment decision consultation.

UPSTREAM: putting it into practice

- Views of UDS as part of standard care, issues with implementing.
- Changes for implementation.

Appendix 7 Quality of urodynamics and urinary flow tests

Urodynamics

In line with the trial monitoring plan (and as specified in the protocol), the quality of UDS testing was reviewed. Compliance with the internationally recognised standard[75] was monitored via a review of ≥ 10% of traces from each centre. The following technical aspects of UDS testing were reviewed:

- Review of unit logs to consider equipment maintenance and calibration testing consistent with manufacturer instructions.
- Measurement of bladder and abdominal pressure, including resting pressures within expected limits.
- Concurrent computing of detrusor pressure.
- Filling to be undertaken within 'physiological rates'.
- Identification of technical problems during studies and suitable remedial action.
- Checks of pressure transmission (e.g. subtraction of cough impulse) during filling and after voiding.
- Trace labelling, such as key events and 'permission to void'.
- Correction for artefacts during computation of BOOI and BCI.
- Correspondence of written report to the original traces.

Urinary flow tests
Centres conducted flow tests according to local practice, the details of which were provided to trial management for review of consistency across all centres. At least 10% of traces from each centre were reviewed.

EME
HS&DR
HTA
PGfAR
PHR

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

Published by the NIHR Journals Library