A MIXED METHODS STUDY TO ASSESS THE FEASIBILITY OF A RANDOMIZED CONTROLLED TRIAL OF INVASIVE URODYNAMIC TESTING VERSUS CLINICAL ASSESSMENT AND NON-INVASIVE TESTS PRIOR TO SURGERY FOR STRESS URINARY INCONTINENCE IN WOMEN: THE INVESTIGATE-I STUDY

Paul Hilton,[1] Natalie Armstrong,[2] Catherine Brennand,[3] Denise Howel,[4] Jing Shen,[4] Andrew Bryant,[4] Douglas G. Tincello,[2] Malcolm G. Lucas,[5] Brian S. Buckley,[6] Christopher R. Chapple,[7] Tara Homer,[4] Luke Vale,[4] Elaine McColl,[3,4] on behalf of the INVESTIGATE studies group.□

Institutional addresses:

[1] Newcastle upon Tyne Hospitals NHS Foundation Trust, UK.

[2] Department of Health Sciences, University of Leicester, Leicester, UK.

[3] Newcastle Clinical Trials Unit, Newcastle University, Newcastle upon Tyne, UK.

[4] Institute of Health & Society, Newcastle University, Newcastle upon Tyne, UK.

[5] Department of Urology, Morriston Hospital, Swansea, UK.

[6] School of Medicine, National University of Ireland, Galway, Ireland.

[7] Department of Urology, The Royal Hallamshire Hospital, Sheffield, UK.

□ Other members of the group are listed in the acknowledgements section.

Email addresses:

Paul Hilton	paul.hilton@ncl.ac.uk
Natalie Armstrong	na144@le.ac.uk
Catherine Brennand	cath.brennand@newcastle.ac.uk
Denise Howel	denise.howel@newcastle.ac.uk
Jing Shen	jing.shen@newcastle.ac.uk
Andrew Bryant	andy.bryant@ncl.ac.uk
Douglas G. Tincello	dgt4@leicester.ac.uk
Malcolm G. Lucas	malcolm.lucas@wales.nhs.uk
Brian S. Buckley	briansbuckley@gmail.com
Christopher R. Chapple	c.r.chapple@sheffield.ac.uk
Tara Homer	tara.homer@newcastle.ac.uk
Luke Vale	luke.vale@ncl.ac.uk
Elaine McColl	elaine.mccoll@newcastle.ac.uk

Corresponding author:

Mr Paul Hilton
c/o Directorate of Women's Services
Level 5, Leazes Wing
Royal Victoria Infirmary
Newcastle upon Tyne
NE1 4LP

Tel: 0191-2825853

Fax: 0191-2825873

Email: paul.hilton@ncl.ac.uk

ABSTRACT

BACKGROUND

The position of invasive urodynamic testing (IUT) in diagnostic pathways for urinary incontinence is unclear, and systematic reviews have called for further trials evaluating clinical utility. The objective of this study was to inform the decision whether to proceed to a definitive randomized trial of IUT compared to clinical assessment with non-invasive tests, prior to surgery in women with stress urinary incontinence (SUI) or stress predominant mixed urinary incontinence (MUI).

METHODS

This was a mixed methods study comprising: a pragmatic multicentre randomized pilot trial; a qualitative face-to face interview study with patients eligible for the trial; an exploratory economic evaluation including value of information study; a survey of clinicians' views about IUT; and qualitative telephone interviews with purposively sampled survey respondents. Only the first and second of these elements are reported here.

Trial participants were randomized to either clinical assessment with non-invasive tests (control arm), or clinical assessment with non-invasive tests plus IUT (intervention arm).

The main outcome measures of these feasibility studies were: confirmation that units can identify and recruit eligible women; acceptability of investigation strategies and data collection tools; and acquisition of outcome data to determine

the sample size for a definitive trial. The primary outcome proposed for a definitive trial was ICIQ-FLUTS (total score) six months after surgery or the start of non-surgical treatment.

RESULTS

Of 284 eligible women, 222 (78%) were recruited; 165/219 (75%) returned questionnaires at baseline and 125/200 (63%) at follow-up. Most women underwent surgery; management plans were changed in 19 (19%) participants following IUT.

Participants interviewed were positive about the trial and associated documentation.

CONCLUSIONS

All elements of a definitive trial were rehearsed. Such a trial would require between 232 and 922 participants, depending on the target difference in primary outcome. We identified possible modifications to our protocol for application in a definitive trial including: clarity over inclusion/exclusions; screening processes; reduction in secondary outcomes; modification to patient questionnaire booklets and bladder diaries. A definitive trial of IUT *versus* clinical assessment prior to surgery for SUI or stress predominant MUI is feasible and remains relevant.

TRIAL REGISTRATION

Current Controlled Trials: ISRCTN 71327395, registered 07/06/2010

Abstract: 342 words

KEYWORDS

Feasibility studies; pilot studies; interview studies; randomized controlled trial; stress urinary incontinence; urodynamics.

BACKGROUND

Urinary incontinence (UI), whilst rarely life-threatening, may seriously influence the physical, psychological and social wellbeing of affected individuals.[1-4] The impact on families and carers may be profound, and the resource implications for health services considerable.[5] Prevalence figures for UI range from 5% to 69% in women 15 years and older, with most studies in the range 25–45%;[6] stress (SUI) or mixed urinary incontinence (MUI) account for 65-85% of cases.[7]

Several methods are used in the assessment of UI to guide management decisions; some of these are non-invasive (*e.g.* urine culture, bladder diaries or frequency volume charts, urine flow rate and post-void residual volume measurement), and some are invasive (*i.e.* require catheterisation). Cystometry, the most commonly used invasive urodynamic test (IUT), looks at the pressure/volume relationships during bladder filling, storage and emptying, with a view to defining a functional diagnosis as distinct from a purely symptomatic one.

The current position of IUT in the diagnostic pathway is not agreed, and practices vary considerably; in a UK survey in 2002 only half of the units surveyed had a guideline on indications for the tests, and 85% carried out cystometry in all women with incontinence.[8] Current guidance from the National Institute for Health and Care Excellence (NICE), however, suggests that cystometry is not required prior to conservative treatments for UI, nor prior to surgery where the diagnosis of SUI is clear on clinical grounds (*i.e.* where there are no symptoms of overactive bladder (OAB) or voiding dysfunction, no anterior compartment prolapse, and no previous surgery for SUI).[9-12]

Changes in available operative techniques, and in particular the introduction of less invasive approaches such as mid-urethral tapes, have resulted in dramatic alterations to surgical practice in recent years.[13] Hospital Episode Statistics (HES) demonstrated a 50% increase in surgery for SUI in the 10 years following the introduction of mid-urethral tapes in 1997, with numbers apparently plateauing at 11,000-13,000 procedures annually in England between 2006-07 and 2012-13.[14] Were the NICE guidance to be applied, the annual savings from more rational use of IUT prior to surgery for SUI, based on 2012/13 national tariff costs (£403 per procedure for Healthcare Resource Group LB42Z)[15] and HES activity data,[14] would be approximately £3.3m. There would also be an additional 'opportunity cost' saving from the alternative use of staff and equipment currently devoted to IUT. On the other hand, it must be recognised that there are increasing concerns about the long-term safety of vaginal mesh implants,[16] which might argue more in favour of increasing use of investigation to ensure the most rational use of surgery.

Two trials looking at the clinical utility of urodynamics in women with SUI have been published recently, both using a non-inferiority design. The VUSIS-1 trial from the Netherlands was terminated prematurely due to slow recruitment after achieving only 23% (59/260) of its planned accrual.[17] In view of the recruitment difficulties with VUSIS-1, the group proceeded to a further study of alternative design, (VUSIS-2) in which all women underwent IUT, and only those with discordant clinical and urodynamic findings were randomized between surgical treatment (as dictated by their clinical assessment) and individualized treatment (dictated by the combination of clinical and urodynamic results); neither

participants nor healthcare professionals involved were blinded to the urodynamic results in either group.[18]

The ValUE trial from the USA defined a non-inferiority margin of 11%;[19] this is equivalent to a standardised difference of <0.8, which may be considered high in statistical terms.[20] A difference in outcome between groups of 11% may also be considered important in clinical terms, potentially influencing the decisions of both clinicians and patients. Notwithstanding these limitations, both studies reported that, in women with uncomplicated SUI, treatment (usually an immediate mid-urethral sling operation) based on basic clinical evaluation is not inferior to individually tailored treatment based on urodynamic findings.

Each of these studies was published during the period of recruitment and follow-up in INVESTIGATE-I.[17, 19, 21] How much they have already influenced clinical opinion and practice, or will do so in the future, is unclear, although a 'point-counterpoint' debate published after these studies makes it clear that there is still a question to be answered.[22, 23] The most recent update of the Cochrane review of urodynamics for the management of urinary incontinence in children and adults included the data from these two trials, yet continued to emphasise the need for larger definitive trials, in which people are randomly allocated to management according to urodynamic findings or to standard management based on history and clinical examination.[24] In addition to NICE[9-12] and the Cochrane Collaboration,[24] the National Institute for Health Research – Health Technology Assessment programme (NIHR-HTA)[25] and the International Consultations on Incontinence (ICI)[26, 27] have also reviewed research literature

on urodynamics, and, along with the James Lind Alliance Urinary Incontinence Priority Setting Partnership,[28, 29] have called for high quality primary research assessing their clinical utility.

But several considerations indicated the need for a pilot trial and feasibility assessment before undertaking a definitive trial. Firstly, to inform sample size calculation. Calculations based upon estimates and assumptions from previously published modelling exercises,[9, 30] and a previous surgical trial[31, 32] are sensitive to parameter values such as the proportion of recruits with SUI,[30] the proportions of poor outcomes in the two arms, and the effect size (target difference) of interest. Calculations based upon data in the most recent Cochrane review of urodynamics indicates that a sample size of over 1600 per arm would be required to address this question.[24] Therefore, given the possible size and cost of a definitive trial, a pilot trial was considered crucial to test assumptions made, give relevant estimates of key parameters, and ensure that a definitive trial would represent value for money from public funds. Secondly, a feasibility assessment could establish whether sufficient clinicians are willing to randomize patients within a definitive trial. IUTs have been widely used in clinical practice over the last 30 years and, despite the lack of evidence of clinical utility, many clinicians look on cystometry as a mandatory part of the investigation of patients with UI, particularly prior to surgical treatment.[33-35] A survey of members of the British Society of Urogynaecology (BSUG) has shown a high level of disagreement with the NICE guidance in this respect,[36] and others have questioned the safety of the recommendations.[37] Finally, a key feasibility objective was to assess patient willingness to participate and identify barriers to and facilitators of participation.

Patients may not so easily see the importance of 'testing a test' in the same way as they might view testing a treatment. Women may be willing to undergo even invasive investigation[38] in the belief that this will inevitably guide them and their clinicians towards appropriate treatment, and away from inappropriate and possibly harmful interventions. In a pilot patient preference study only 32% of women were prepared to be randomized.[38]

Recognising that a pilot randomized controlled trial (RCT) alone was probably inadequate to address the complexities of feasibility for a definitive trial in this aspect of healthcare, the INVESTIGATE-I study comprised an external pilot RCT, an exploratory health economic analysis and value of information study, a national survey of relevant clinicians, and separate qualitative interview studies with patients eligible for the trial and clinicians responding to the survey. Only the first and second of these elements are reported here.

The original study protocol was published in this journal;[39] two later amendments were approved by the Research Ethics Committee, and the final version of the protocol (v1.2) is available on the NIHR website http://www.nets.nihr.ac.uk/projects/hta/0922136. The clinician survey and interview study have been published in full previously,[40, 41] and a separate publication is planned for the economic evaluation and value of information study.[42] This report therefore, whilst drawing conclusions from the whole collection of studies, focuses on the pilot trial itself, and the qualitative interview study with trial participants.

METHODS

The conduct of this study was in accordance with the ethical principles set out in the Declaration of Helsinki (2008) and the Research Governance Framework for Health and Social Care (second edition, 2005).[43] Application for ethical approval was made through the Integrated Research Application System (IRAS), and a letter of favourable ethical opinion was obtained from Newcastle & North Tyneside 1 Research Ethics Committee on 6th January 2011 – reference no. 10/H0906/76. All elements of the study were approved by local Research and Development offices at Newcastle upon Tyne Hospitals NHS Foundation Trust (28/03/2011), Gateshead Health NHS Foundation Trust (29/03/2011), Abertawe Bro Morgannwg University Health Board (23/06/2011), Sheffield Teaching Hospitals NHS Foundation Trust (07/07/2011), Northumbria Healthcare NHS Foundation Trust (25/07/2011), University Hospitals of Leicester NHS Trust (09/08/2011), City Hospitals Sunderland NHS Foundation Trust (30/05/2012), South Tees Hospitals NHS Foundation Trust (09/07/2012) and South Tyneside NHS Foundation Trust (17/09/2012); hence the favourable ethical opinion was applicable to all NHS sites taking part in the study.

The objective of the feasibility study (INVESTIGATE-I) was to inform the decision as to whether to proceed to a definitive RCT of the clinical and cost-effectiveness of IUT compared to basic clinical assessment with non-invasive testing in women potentially suitable for surgical treatment of SUI or stress predominant MUI, and whether any refinements to the proposed definitive trial design were warranted.[44-48]

1. PRAGMATIC MULTICENTRE RANDOMIZED PILOT TRIAL.

The pilot RCT was designed to rehearse the methods and processes of any future definitive RCT.

UNITS RECRUITING TO THE TRIAL

Recruitment to the pilot trial was initially limited to six specified units; these were a mix of specialist urogynaecology (Newcastle upon Tyne and Leicester) and female urology (Sheffield and Swansea) departments in university teaching hospitals, providing secondary and tertiary level care, and general gynaecology units in district general hospitals, providing secondary care services (Wansbeck Hospital, Northumberland, and Queen Elizabeth Hospital, Gateshead).

In order to improve adherence with recruitment targets, and to test the processes for possible future use, two Patient Identification Centre (PIC) sites (Sunderland Royal Hospital and South Tyneside District General Hospital) and one additional full recruiting site (South Tees Hospitals NHS Foundation Trust) were added during 2012.

INCLUSION AND EXCLUSION CRITERIA

Inclusion criteria

Inclusion criteria for the pilot RCT (and anticipated inclusion criteria for any future definitive RCT) were:

- Clinical diagnosis of SUI or stress predominant MUI.
- Women must state that their family is complete.
- Women should have undergone a course of pelvic floor muscle training (± other non-surgical treatments for their urge symptoms) with inadequate resolution of their symptoms.
- Both the woman herself and her treating clinician should agree that surgery is an appropriate and acceptable next line of treatment.

Exclusion criteria

Exclusion criteria for the pilot RCT (and anticipated exclusion criteria for any future definitive RCT) were:

- Symptomatic utero-vaginal prolapse requiring treatment.
- Previous surgery for urinary incontinence or pelvic organ prolapse.
- Urodynamic investigation within the last three years.
- Neurological disease causing urinary incontinence.
- Current involvement in competing research studies, e.g. studies of investigation or treatment of urinary incontinence.
- Unable or unwilling to give competent informed consent.

RECRUITMENT

Potential trial recruits were identified by research nurses prior to attending new or follow-up appointments for SUI or MUI. A short Patient Information Leaflet (PIL) was included with a letter of invitation, with new appointments or reminder

letters for follow-up appointments. A full (6 page) PIL was provided on request. The study information was discussed at the first hospital visit; women declining to take part underwent further investigation and or treatment as clinically appropriate at the same visit. Written consent was obtained from those agreeing to take part, before randomisation. To ensure concealment of allocation, randomization was undertaken by an internet-accessed computer randomization system held by the Newcastle Clinical Trials Unit (NCTU); randomization between intervention and control was 1:1, and was stratified by centre using random block length. It was neither feasible nor appropriate to blind participants or clinicians (investigating and operating) as to the allocation of investigation strategy.

SAMPLE SIZE

The sample size for the external pilot trial was determined pragmatically, using the recommended minimum of 30 participants per arm.[47] It was hoped that 60 would be retained per trial arm to investigate the distribution and key parameters of the outcome measures. Previous trials in the area of pelvic floor dysfunction, including investigation,[49] surgical,[32, 50, 51] and non-surgical treatments[52] suggested average attrition rates of 13% (7-20%) between identification and randomization, 16% (6-20%) between randomization and treatment, and 13% (9-20%) between treatment and follow-up at six months. Based upon the more pessimistic figure in each case, it was estimated that a total of 240 eligible patients should be approached, allowing for a 50% overall attrition.

INTERVENTIONS

Patients were randomized to receive either:

- 'no IUT' - basic clinical assessment supplemented by non-invasive tests as directed by the clinician; these included frequency/volume charting or bladder diary, mid-stream urine culture, urine flow rate and residual urine volume measurement (by ultrasound), or
- 'IUT' - basic clinical and non-invasive tests as above, plus invasive urodynamic testing (IUT). Dual-channel subtracted cystometry with simultaneous pressure/flow voiding studies is the most commonly applied technique in the evaluation of patients prior to surgery for SUI in most centres; videourodynamics and ambulatory bladder pressure monitoring were also permissible at the discretion of the clinician.

Further investigation was undertaken where appropriate at the same visit or a later one, as per local practice, and the treatment plan formulated.

OUTCOME MEASURES

The collection of the outcome measures for a future definitive RCT was piloted, to assess data yield (*e.g.* percentage of recruited participants returning completed questionnaires) and quality (*e.g.* completeness and consistency of responses within returned questionnaires). This information was collected to guide the choice and mode of administration of questionnaires and data collection tools in any future definitive RCT.

The primary outcome rehearsed in the pilot RCT was a patient reported outcome measure (PROM):

- The combined symptom score of the International Consultation on Incontinence female lower urinary tract symptoms questionnaire (ICIQ-FLUTS) at six months after treatment.[31]

Secondary outcomes rehearsed were:

- General health questionnaire (SF-12v2™ Health Survey © 1994, 2002 by QualityMetric Incorporated and Medical Outcomes Trust)[53], and EQ-5D-3L © 1990 by EurQol Group [54])
- Quantification of urinary leakage (three day bladder diary, and ICIQ-UI SF)[55]
- Prevalence of symptomatic '*de novo*' functional abnormalities including voiding dysfunction and detrusor overactivity (using subscales in ICIQ-FLUTS,[31] with cystometric investigation in symptomatic patients)
- The impact of urinary symptoms on quality of life (ICIQ-LUTSqol and UDI);[56, 57] the latter measure was included since it was used in the VUSIS and VALUE trials.[18,19]
- Use of health services and costs to the NHS and to patients

BASELINE ASSESSMENT OF STUDY OUTCOMES

Following consent and randomization, patients were given a pack of baseline study outcome questionnaires. Participants were asked to complete the

questionnaires at home, within two weeks of receipt, and post them to the central trial office using a prepaid envelope.

SUBSEQUENT TREATMENT WITHIN THE TRIAL

Following investigation, it was expected that women randomized to the 'no IUT' arm of the study would undergo surgical treatment. The choice of operation was left to the individual surgeon and woman; since only primary cases were included, it was anticipated that in most cases this would be either a retropubic or transobturator foramen mid-urethral tape procedure. It was expected that those randomized to the intervention 'IUT' arm would have similar surgical treatment when urodynamic stress incontinence (USI) was confirmed. Where other diagnoses were identified following investigation, alternative treatments might be offered, informed by which other conservative treatments had previously been tried. These included bladder retraining, anti-muscarinic drug treatments, neuromodulation, botulinum toxin injections (where detrusor overactivity (DO) was diagnosed), or clean intermittent self-catheterisation (where a voiding dysfunction was identified). In all centres the treatment algorithm employed was in keeping with the then current NICE recommendations (2006).[9]

FOLLOW-UP

Clinicians arranged post-operative follow-up or other outpatient review, as per their normal practice and timing. Women were sent a pack of follow-up study

outcome questionnaires and bladder diaries along with a prepaid envelope at six months after surgery, or the start of any non-surgical intervention, or period of 'watchful waiting'. They were asked to complete and then post them to the central trial office. Those failing to return questionnaires within one month were contacted by a research nurse by telephone, to encourage responses. In the last nine months of the study the option of completing the questionnaire over the telephone with the research nurse was also given to participants during the reminder telephone call. Those who did not return the questionnaires after a telephone reminder were sent a second copy of the questionnaires. Each patient's withdrawal or completion of study follow-up was documented in the case report form (CRF).

2. QUALITATIVE INTERVIEWS WITH WOMEN ELIGIBLE FOR THE PILOT TRIAL.

Interviews were carried out to explore women's understandings and experiences of the study, including the consent processes and their decision to participate. Purposive sampling was used to invite women from a range of ages, trial participation status (randomized and retained to final follow-up; randomized but did not provide full follow-up data), allocation status (IUT or basic assessment), treatment received (surgery or conservative management), and study site. It was also intended that women who declined randomization would be interviewed.

Women were approached at the end of the trial, so as to capture both their reasons for agreeing to participate and their overall experience of taking part in the study. A specific Participant Information Leaflet was provided for the

interview study and written consent was obtained from all interviewees. The interviews were carried out face-to-face by an expert qualitative interviewer (see acknowledgements) and were audio-recorded and transcribed verbatim.

The interviews were semi-structured, using a prompt guide with broad topic areas, but the emphasis was on encouraging women to discuss their own perspectives freely and allowing them to raise issues that were important to them. The interviewer prompted as appropriate to ensure that all views were fully explained, and the meaning of participants' responses clear. The prompt guide was developed from a literature review and discussions within the project team and was modified as the interviews progressed to incorporate issues raised by earlier interviewees.

Analysis took place alongside data collection which continued until saturation of themes was reached and interviews no longer generated new concepts. All completed interviews were included in the analysis. Analysis was based on the constant comparative method[58], and aided by *NVivo 10* software (© QSR International, Warrington, UK). Data analysis was carried out by an experienced qualitative researcher (see acknowledgements) under the supervision of NA. To maximise the credibility and rigour of the analysis, NA regularly reviewed the coding scheme and interview transcripts and any differences in interpretation were discussed and agreed. Further details of the methods are published in full in the protocol document.[39, 59]

SYNTHESIS OF FINDINGS

The analytic framework proposed by Bugge et al (2013) was used to summarize findings from the pilot trial and participant interviews;[45] this framework comprises 14 methodological issues, derived from the work of Shanyinde et al (2011) on what needs to be evaluated in pilot and feasibility studies.[60]

This analysis is followed by the 3-step ADePT process, involving:

i) Deciding on the type of problem experienced (Type A – the issue is likely to be a problem only for the trial; Type B – the issue is likely to be a problem for both the trial and the real world; Type C – the issue is likely to be a problem only for the real world), and the associated evidence;

ii) Identifying the range of possible solutions and the evidence to support those solutions, including assessment of the potential effectiveness and potential feasibility of each option;

iii) Assessing the best options.

RESULTS

The summary of methodological issues,[60] and their analysis after Bugge et al, 2013,[45] is given in *Table 1*.

PRAGMATIC MULTICENTRE RANDOMIZED PILOT TRIAL

SCREENING, RECRUITMENT AND RANDOMIZATION

The screening, recruitment, randomization and trial follow-up are summarised in the CONSORT diagram shown as *Figure 1*. Overall, 771 women were identified and were sent the patient information sheets. Of those 284 were deemed eligible for the trial, (37% screen positive). The reasons for non-eligibility, which varied between centres, are shown below in *Table 2*. One centre accounted for more than half the women screened (399; 52%).

Of the 284 women screened positive, 222 agreed to randomization into the trial, giving a trial consent rate of 78%. This recruitment total (222) represented 93% of the planned sample size (240) for the pilot trial. Overall, 110 women were randomized to the 'no IUT' arm and 112 to the 'IUT' arm. Immediately after randomization it became apparent that one woman in the 'no IUT' arm was ineligible for the trial and she was withdrawn leaving a total of 221 eligible patients randomized (109 in the 'no IUT' arm and 112 in the 'IUT' arm).

Monthly recruitment is shown in *Figure 2*. Regulatory requirements took approximately three months longer than anticipated, and recruitment targets were revised accordingly. The rate of accrual over time was significantly less than

required; several steps were introduced to improve recruitment, including the incorporation of additional clinicians at two of the existing sites, and the establishment of an additional full recruiting site and two Participant Identification Centre (PIC) sites; a nine-month unfunded extension to the recruitment period was agreed with the study funder. Newsletters reporting the progress of the pilot RCT and regular recruitment updates were provided to clinicians in order to maintain their engagement.

The number of participants recruited per recruiting month (*i.e.* between the completion of all site specific regulatory requirements and the end of the study) varied between 0.4 and 3.9 per month at the original sites (mean 1.9); at the additional full recruiting site this figure was 2.5 per month; the PICs did not identify any potentially eligible patients for referral to a recruiting site in the eight months that they were active.

Table 3 provides the demographic data by trial arm; the consistency of these variables between 'IUT' and 'no IUT' arms confirms the validity of the randomization process.

RETENTION

Two women in the 'IUT' group withdrew because they were unhappy with their allocation. Baseline questionnaires were sent to 219 women and returned by 165 (a 75% response rate overall, 72% 'IUT' arm and 79% 'no IUT' arm). At six months follow-up, questionnaires were returned by 63% (125/200), (56% (54/97) 'IUT' arm and 69% (71/103) 'no IUT' arm).

COMPLETENESS OF DATA COLLECTION

Not all women fully completed each questionnaire although missing values within individual scales were few. The columns to the right-hand side of *Table 4* show the proportion of each questionnaire or subscale that could be calculated from the data provided.

COMPARISON OF RESPONDERS AND NON-RESPONDERS TO SIX-MONTH QUESTIONNAIRE

Given the high rate of non-response to the six-month questionnaires, a comparison of responders and non-responders was made on the basis of their clinical follow-up. A total of 135 women had a postoperative follow-up visit documented on the study database; 93 actually attended an outpatient clinic, and 42 had a review by telephone (routine practice in three of the centres).

Of the 125 women who returned follow-up questionnaires at six months after treatment, 83 had clinical follow-up, of whom 12/83 (14.5%) described bothersome urinary symptoms, and 9/83 (10.8%) had clinically significant examination findings. Of the 81 who failed to return follow-up questionnaires at six months, 52 had clinical follow-up, of whom 5/52 (9.6%) described significant urinary symptoms, and 4/52 (7.7%) had clinically significant examination findings.

Whilst those women returning the six-month questionnaires somewhat more often had bothersome symptoms or clinically significant examination findings at clinical review than those failing to do so, the numbers do not allow meaningful statistical comparison.

QUESTIONNAIRE DATA

Baseline

Table 4 shows the distribution of the questionnaire scales at baseline by trial arm. The distribution of ICIQ-FLUTS total score at baseline was fairly symmetrical with a mean of 16.9 (SD 5.7) in the 'IUT' arm and 16.4 (SD 6.3) in the 'no IUT' arm. The distributions of the other scales and subscales were similarly well matched between the 'IUT' and 'no IUT' arms and were fairly symmetrical.

Six-month follow-up

Table 4 also shows the distribution of the questionnaire scales at six-month follow-up by trial arm. For all scales, typical scores were much lower than at baseline. It is difficult to interpret any difference in mean scores between baseline and six-month follow-up from *Table 4*, because of the small sample size and the number of women who provided baseline data but for whom no six-month questionnaire data are available. *Table 5* shows the distribution of the paired changes in scale scores for those women who had completed both questionnaires. It can be seen that the mean change in ICIQ-FLUTS total score was 7.8 in the 'IUT' arm and 9.3 in the 'no IUT' arm. Typically, there was a marked drop in these scores over six months, but little difference in the mean changes between the trial arms; this pattern was also seen in the other four scales, although no formal comparison between arms is appropriate in a pilot study.

TREATMENT DATA

In the 'IUT' arm, 82 women (80%) received surgery, compared to 103 (95%) in the 'no IUT' arm. The distributions of operation type, grade of surgeon, anaesthetic technique and use of antibiotic prophylaxis were similar between the trial arms.

One woman in the 'no IUT' arm and four (4%) in the 'IUT' arm decided to defer any treatment initially (designated as '*watchful waiting*'). A further 15 women (15%) in the 'IUT' arm underwent lifestyle changes or other non-surgical treatments. As routine in continence management, more than one lifestyle change was commonly documented, and other non-surgical treatments were often used in combination; 28 treatments were applied in these 15 women. Despite prior (unsuccessful) completion of a course of supervised pelvic floor muscle training (PFMT) being an inclusion criterion for the trial, six women underwent further PFMT alone (n=2) or in combination with other non-surgical treatments (n=4).

ADVERSE AND SERIOUS ADVERSE EVENTS

Only two serious adverse events were reported. One woman in the 'IUT' arm experienced bleeding from sub-urethral incision 12 days after surgery and one woman in the control arm was treated for breast cancer by mastectomy shortly after her surgery within the trial; whilst the first clearly related to the incontinence treatment, neither event was categorised as being related to the trial intervention (IUT).

In addition, 23 adverse events were reported in 22 women; these included three operative bladder injuries (3/185=1.6% perforation rate) and two vaginal injuries. Six episodes of urinary tract infection (UTI) were reported, two in the 'IUT' arm,

and four in the 'no IUT' arm; all occurred following surgery, and none immediately after IUT.

CALCULATION OF POTENTIAL SAMPLE SIZE OF DEFINITIVE TRIAL

Based upon the trial results, the study team decided that differences of 2, 3 or 4 units on the ICIQ-FLUTS scale would be realistic and potentially clinically important differences that might be achieved.

Given these estimates of effect size, a standard deviation of 7 for paired changes between baseline and follow-up, Type I error of 5% and Type 2 error of 10%, total sample size estimates for any definitive trial fall between approximately 200 and 900 women recruited (*Table 6*). These estimates are considerably less than calculations based upon data in the most recent Cochrane review of urodynamics, which indicate that a sample size of over 1600 per arm would be required to address this question.[24] With a recruitment rate of 78%, recruitment of between 200 and 900 would require between approximately 300 and 1200 eligible women to be approached; in turn, with a screen positive rate of 37%, this would mean between approximately 800 and 3000 women would need to be identified for screening for eligibility; these ranges depend upon the effect size.

PATIENT INTERVIEW STUDY

All 59 eligible women who declined to participate in the pilot trial were invited to interview but none was willing.

A diverse sample of 111 pilot trial participants was invited to take part in the interview study, including participants from different study sites, the two study arms, a wide range of ages, and those who did and did not complete all follow-up. A total of 36 women indicated they were willing to be interviewed, but of these two withdrew from the interview study before the interview could be arranged, and another had moved and so was no longer covered by our research governance approvals. Of the remaining 33, 29 were interviewed before saturation of themes was reached and the last four were not interviewed as they were from groups already well represented in the sample. Interviewees were between 35 and 75 years of age, came from five of the seven full trial centres, and included participants from both 'IUT' (16) and 'no IUT' (13) arms.

THE INVITATION TO PARTICIPATE, AND REASONS FOR AGREEING

Women's first reactions to receiving the invitation to participate in the pilot study were almost exclusively positive. The decision to take part was commonly made quickly and easily, and very few reported feeling the need to talk with family or friends as part of the decision-making process.

> *WAS IT AN EASY DECISION TO MAKE?*
>
> *Yes, very.*
>
> *DID YOU MAKE IT ON YOUR OWN?*
>
> *Yes, (Participant 10)*

As is commonly found in other studies,[61-63] many women's reasons for participation were altruistic and included wanting to help research, to help others

with the same condition, and to make some form of repayment for the help and treatment they were receiving.

Participating in the pilot did not seem to require a lot from them and so no particular participation burden was perceived.

> *She explained it very clearly and said all it is basically is just to monitor how many times you go to the toilet, and how much you drink, and roughly how much your output was. And to me I thought that wasn't a big problem. Only a few minutes of your time in your day, just to keep track. (Participant 04)*

THE INFORMATION PROVIDED ABOUT THE STUDY

Reactions to the written information were mostly positive – it was regarded as clear and informative and there was enough information for women to be able to make a decision about taking part. The short version was sufficient for some and the flow diagram was popular. Others liked to have the fuller detail in the longer version. Overall, most people found it helpful, describing it as easy to read, informative, and pitched at the right level.

> *So everything was really well explained you know, so yeah I mean I can't fault it really, no I was well impressed with it all. (Participant 25)*

The use participants made of the material varied – some read it once only or just skimmed it, others read it more than once and a small number did additional research about the study on the internet.

I think I just read it, I didn't take too much in I think, I think I was just so looking forward to getting my operation that is all I was really erm... really bothered about. I don't think I read too much about the ins and outs of the study. (Participant 20)

Basically I just went on-line and looked at the various things and just erm... just looked at the study. (Participant 15)

Some were happy with the verbal information at the time of their consultation and paid little attention to the written material, particularly the longer version.

Personally I wouldn't bother with the big one, I think that there is enough information, and if you get good medical staff to start with like I did, who actually took the time to go through it with you and say this is what this says, now read it on there, erm... so I think if you get that then you certainly don't need the bigger one. (Participant 07)

Suggestions for how the information might be improved were limited but included keeping it as short and concise as possible and distributing prior to the consultation as some women reported feeling anxious at the consultation and did not initially pay much attention to the information. Given that some women valued the verbal information they received from clinical staff more than the written information, being able to go to the consultation with questions prepared may have been helpful.

UNDERSTANDING OF THE STUDY

Participants' understanding of the study was broadly good, although there were some cases in which people appeared confused about the overall aim. Overall,

there was a generally good understanding that the study was assessing the value of a particular diagnostic test rather than the treatment they would ultimately receive. Many talked explicitly about how, while participation in the study could influence the route they took to treatment, it was ultimately unlikely to change the final outcome. Establishing this was often important to securing their participation.

> *I remember asking him "so if I don't have the test will it have any effect on any treatment I have, and will it have any effect on you deciding what I need?" No he said, it was purely for this investigation. (Participant 22)*

Not all participants understood the study in this way, though. A small number, when asked to explain what they thought the study was about, did focus on the subsequent treatment rather than the invasive testing.

> *I think it's about finding the right appropriate erm... ways forward to treat people with urinary problems. Erm... whether surgery or invasive treatment is appropriate or whether there is another kind of treatment that might be more beneficial. (Participant 17)*

The principle of random allocation to one of two possible groups was generally well understood. There were, however, a small number who thought that participation in the study automatically meant they would avoid the invasive tests.

> *DID YOU THINK THERE WAS A POSSIBILITY THAT YOU MIGHT HAVE THE INVASIVE TESTS?*
>
> *Erm...no I think the registrar said to me if I signed up for the study I wouldn't have them. (Participant 08)*

EXPERIENCES OF STUDY PARTICIPATION

The first set of questionnaires participants were asked to complete at baseline was generally described as simple to fill in, easy to understand, and straightforward.

> *HOW DID YOU FIND THE QUESTIONNAIRES YOU WERE ASKED TO COMPLETE AT THE BEGINNING?*
>
> *Simple.*
>
> *WERE THEY TIME CONSUMING AT ALL?*
>
> *No not particularly. (Participant 01)*

A few minor issues were raised: there wasn't always a box to tick that was applicable to them; some questions were hard to answer (e.g. when asked to work out costs or where judgement was called for); and some thought the questions were a little repetitive.

> *Sometimes there wasn't, you know how there were tick boxes kind of thing, it...none of those were really the answer that I wanted to give. (Participant 11)*

There were also some comments on the practical challenges associated with measuring urine output for the bladder diary.

> *I found it more difficult to collect the urine. You know to get down to it and have clear, clear days to get on with it. (Participant 18)*

The second set of questionnaires sent out six months after treatment were similarly felt to be relatively simple to complete. However, given that many had

had successful treatment and now had few, if any, symptoms to report, the questions did not always seem relevant. Indeed, one participant reported having called the study office to check she had been sent the right questionnaires, and others were a little concerned it might appear that they had not completed the questionnaires at all because so much was not now applicable to them.

> *I actually sent it back with absolutely nothing on it at all because it said "have you been to visit the doctor in 6 months", and I hadn't and it said go to the next section, and go to the next section and so by the end of it, there was nothing on it and I sent it back completely blank and I thought they will think I have not bothered filling this in. (Participant 14)*

While some actually found completing the six-month questionnaires quite enjoyable (as it underlined for them how successful the treatment had been), others reported finding them burdensome and irrelevant now they had few or no symptoms to report.

> *Not relevant at all, not to me anyway. Yes, because I mean the problem was solved then so, why harp on about how many pads am I wearing now because I don't wear them, simple as that, nothing. (Participant 09)*

This seemed particularly to apply to the bladder diaries.

> *It did want another bladder diary I think afterwards and I have not completed the bladder diary because I just didn't get round to it to be honest with you. I had it in my bag to take to work with me and I just didn't get round to doing it. (Participant 21)*

DISCUSSION

The findings and implications of this pilot are considered in subsequent sections across a number of aspects of the trial design.[60] In terms of the ADePT approach, the problems identified related to aspects of trial process and were therefore classified as Type A – issues likely to be a problem only for a trial, but not in the real world.[45]

Overall, the logistics and study procedures were seen to be adequate and functional in most areas, and important insights were gained to inform the design and efficient conduct of any future definitive trial. These include: allowing a realistic time frame for regulatory approval and site start-up; clarity over inclusion/exclusions; modifying screening processes; reduction in secondary outcomes; modification to patient questionnaire booklets and bladder diaries; and employing a range of strategies to retain trial centre engagement (*e.g.* website, newsletters, recruitment updates).

ELIGIBILITY, RECRUITMENT, CONSENT AND RANDOMIZATION

We found that 37% of women screened were deemed eligible for the trial. This figure varied between centres, as did the declared reasons for ineligibility. More than half of all the women screened were from one centre. It is likely that the assiduousness of recruiters and interpretation of eligibility criteria differed between centres. Running screening training exercises might be considered for a future definitive trial to ensure similar screening standards and practices and an 'assumed eligibility' approach in all centres. This should be feasible *e.g.* by

'clustering' centres geographically and carrying out training exercises alongside site setup visits; we do not however have evidence of the effectiveness of this proposed solution.

Recruitment was initially slow, and was more successful in some centres than others. Recruitment was initially delayed by the fact that ethical and regulatory requirements for a multi-centre study took longer than expected, and any definitive trial should determine and allow a realistic timeframe for this.

Once approvals were in place, it was necessary to expand the number of planned centres and clinicians within centres to meet recruitment targets; this highlights the need for rigorous and realistic site feasibility assessments prior to site selection and setting and on-going monitoring of individual site targets.

Whilst there is little high quality evidence to support their use,[64] a range of strategies was used to retain trial centre engagement such as regular recruitment updates and newsletters. However we were eventually able to recruit patients from all our study centres in sufficient numbers to confirm that recruitment was feasible.

Of those women who screened positive, 78% consented to enter the trial. Data from the patient interviews suggested that most women reacted positively to the invitation to take part, and found the information provided about the study to be clear. There was no clear preference for either the shorter or longer version of the patient information sheet. The principle of random allocation to one of two trial arms was generally well understood by participants. The randomization

procedure led to similar sized groups that were well balanced on baseline variables.

COMPLIANCE WITH AND ACCEPTABILITY OF INTERVENTION

Most patients received the 'IUT' (91%) or 'no IUT' group tests (99%) to which they were allocated. However, two patients withdrew from the trial because they were unhappy to be randomized to the 'IUT' arm, one failed to attend the appointment for IUT, and four other patients in the IUT arm did not undergo invasive tests for unspecified reasons.

OUTCOME ASSESSMENT, SELECTION OF MOST APPROPRIATE OUTCOMES AND PARTICIPANT RETENTION

Completion rates were relatively high for all questionnaires, and they had a similar rate and spread of missing items. Rates of loss to follow-up after treatment were significant, however, and whilst 75% of women had either face-to-face or telephone follow-up (typically at two to three months) after surgical treatment, only 56% (63% of those circulated) returned follow-up questionnaires at six months.

It is recognised that the completion of questionnaires can be burdensome for participants,[65] and this may be particularly the case for those with few or no symptoms. We found some evidence in the patient interview study to suggest that women were less likely to return questionnaires if they were satisfied with the results of their treatment, which may account for the number of blank questionnaires returned at six months.

In any future definitive trial it would be necessary to ensure a higher questionnaire response rate. The UDI was the fourth instrument in a booklet of six questionnaires in total, and had a slightly lower completion rate at both baseline and six months. The questions in ICIQ-UI SF overlap considerably with those in the longer ICIQ-FLUTS and so we recommend omitting both UDI and ICIQ-UI SF from any definitive trial to reduce respondent burden. We anticipate that this may improve completion of the remaining items. Greater emphasis needs to be placed on the importance of returning a completed questionnaire even in the absence of any remaining symptoms. Alternative modes of completion for follow-up questionnaires (*e.g.* telephone or web based), and providing incentives to return questionnaires, are further evidence-based strategies that might enhance retention rates for data collection.[66, 67]

Bladder diary data and pad test use were poorly completed in our pilot. This may be because many of the women would have completed similar diaries or frequency/volume charts earlier in their continence assessment; it may be seen as rather more intrusive than simple questionnaire responses; it is possible that the diary design resulted in inconsistent completion of pad-use data. The trial recruitment process enrolled only women with SUI or stress-predominant MUI, and the diary data did not show any evidence of abnormal urinary frequency or nocturia and there appeared to be no change at six months in either arm (other than in pad-use). In order to increase the completion rate of incontinence episode data, diary data and pad use might be omitted or modified in any definitive trial.

Alternative modes of completion for follow-up questionnaires, such as by telephone or online, and providing modest incentives to return questionnaires,[66, 67] are further evidence-based strategies that might enhance retention rates for data collection. A further possibility is to link questionnaire completion at follow up to the face-to-face clinic review, thereby allowing a check by a research nurse or trial coordinator of item completion before patients leave the clinic area; however, this would have required a change to the current practice of some units, and risked the pragmatic nature of the trial.

SAMPLE SIZE CALCULATION FOR A DEFINITIVE TRIAL

Sample size estimates were calculated for target differences of 2,3, and 4 units in ICIQ-FLUTS, using the standard deviation of the primary outcome data from the pilot trial. However, a monograph on ways of specifying a target difference for a trial recommended that estimates of sample size should be determined by more than one approach.[68] In any definitive trial, the following data sources might be amongst those considered:

1. Clinician opinion
2. Data from the external pilot trial
3. A value of information study (not included here, but forming part of a separate report)[42]

A survey update in June 2013 of consultant members of BSUG and BAUS-SFNUU sought their views on what constitutes a minimum clinically significant target difference in ICIQ-FLUTS combined score. However, the ICIQ-FLUTS scale has not been used in many published studies to date, and, perhaps because it is therefore

not familiar, only 50% of consultants responding expressed an opinion. There was no evidence of a common opinion: given a choice of seven ranges of the scale to define a clinically important difference (from 1-4 to >24), all these ranges were chosen by at least one clinician, with the modal range being 9-12. In separate discussions, members of the study team did not find it easy to choose a target difference based on the limited use of the scale so far. The current lack of data from published trials using ICIQ-FLUTS, and therefore evidence on which to base expert judgement, casts some doubt of the usefulness of a survey of experts in this situation.

When the pilot trial results became available, it was apparent that the distribution of the ICIQ-FLUTS total score at six months, and the difference between scores at baseline and six months, typically had low values. The mean score (SD) at six months in the 'no-IUT' arm was 6.9 (5.0) and the mean change between baseline and six months was 9.3 (7.3). It was apparent, therefore, that it is not realistic to expect differences in mean outcomes between trial arms in the order of 9-12 units, as proposed in clinician survey responses. Based upon the trial results, the study team decided that differences of 2, 3 or 4 units would be realistic differences that might be achieved in any comparison of an intervention for women eligible for a future trial.

Given the observed standard deviations, these target differences of 2, 3 or 4 units are equivalent to standardised effect sizes of 0.29, 0.43 and 0.57 when comparing mean changes in score over six months. In contrast, a difference of 9-12 units would equate to a standardised effect size of 1.5-2, which is a very large

difference; many trials are planned on a standardised effect size of around 0.5. Cohen has suggested that standardised differences of 0.2, 0.5 and 0.8 correspond to 'small', 'medium' and 'large' effect sizes.[20]

If a study is planned on the basis of a 'realistic' value for the target difference, then consideration has to be made of whether this is also a 'clinically important' difference. If it is clear that this is not a 'clinically important' difference, then there are real doubts as to whether the trial should take place. It was felt that a difference of around three units would also be of clinical interest since a decrease of this level would equate to complete recovery for one of the symptoms assessed in the ICIQ-FLUTS score.

In this pilot trial we identified 771 women for screening from seven centres over the course of 114 centre screening months (approximately 6.8 women/centre/screening month). Extrapolation of these figures would require 120-480 centre screening months to achieve the recruitment of 200-900 women. This would mean 4-20 centres recruiting for approximately 30 months or 6-30 centres recruiting over 18 months.

CONCLUSIONS

Overall, the pilot trial can be considered a success and a definitive trial is feasible and remains necessary. The study procedures were seen to be adequate and functional in most areas, and important insights were gained to inform the design and efficient conduct of a future definitive trial.

Lessons were learned in how to manage the time needed to bring multiple centres online through the UK regulatory process; variation in recruitment likely from multiple centres has been observed and the importance of standardised and assiduous screening recognised; effective methods of communication to keep staff engaged through the lifetime of a long study have been rehearsed and refined. Refinements in the data collection process that will improve the quantity and quality of the data for a definitive trial have been identified.

Although recruitment was initially slow, patients were recruited from all study centres in sufficient numbers to confirm that recruitment is feasible, and that women are happy to engage with the study objectives and be randomized. Participants were very positive about the study, and in particular allayed fears over whether research to 'test a test' would be seen as important. The interviews also offered suggestions as to how the experience of participation could be improved and data collection maximised.

Based upon a range of target differences derived from the observed clinical outcomes in this pilot RCT, any definitive trial may need to recruit between 200 and 900 women. With recruitment rates also based upon the pilot RCT, this

would mean 4-20 centres recruiting for approximately 30 months or 6-30 centres recruiting over 18 months.

LIST OF ABBREVIATIONS

BAUS-SFNUU – British Association of Urological Surgeons – Section of Female, Neurological, and Urodynamic Urology; BSUG – British Society of Urogynaecology; CLRN - Comprehensive Local Research Network; CRF – case report form; DO – detrusor overactivity; HES – Hospital Episode Statistics; ICER - incremental cost-effectiveness ratio; ICIQ-FLUTS=International Consultation on Incontinence modular questionnaires: Female Lower Urinary Tract Symptoms questionnaire; ICIQ-LUTSqol=ICIQ Lower Urinary Tract Symptoms quality of life questionnaire; ICIQ-UI SF=ICIQ Urinary Incontinence Short Form questionnaire; IUT – invasive urodynamic testing; MUI – mixed urinary incontinence; NIHR - National Institute for Health Research; OAB - overactive bladder; PCQ – participant costs questionnaire; PFMT - pelvic floor muscle training; PIL – Patient Information Leaflet; QALY - quality of life year; RCT - randomized controlled trial; SUI – stress urinary incontinence; UDI=Urogenital Distress Inventory; UI – urinary incontinence

COMPETING INTERESTS

The authors have no current personal financial interests; the following non-financial interests are declared:

PH: Previous chair of NICE Guideline Development Group (GDG) on urinary incontinence (UI) in women (2004-'07); member of faculty of International Consultations on Incontinence – Urodynamics committee (1997-'99 & 2000-'02); Clinical assessment committee (2003-'05); Surgery in women committee (2007-'09); James Lind Alliance working partnership on research priorities in urinary incontinence (2007-'09); previous commercial research funding for trials of surgery for stress incontinence from *Gynecare* (1998–2003) and *Gyne Ideas* (now *Mpathy Medical*; 2001–'03).

DGT: Has received grant income for investigator-initiated research from *Ethicon*; paid consultancies from *Ethicon* and *Allergan* for teaching and advisory boards; financial support to attend conferences in 2012/13 from *Ethicon* and *Astellas*. All income managed by the University of Leicester Research and Business Development.

MGL: Research funding received into department from Astellas, Allergan, Pfizer, Astratech, Urovalve Inc. Member of NICE Guideline development groups for incontinence in women (2006) and Male LUTS (2010). Chairman of European Association of Urology Guideline Panel on Incontinence 2009 – 2013.

EM: Is a NIHR Journals Editor.

NA, CB, DH, JS, AB, BSB, CRC, TH, and LV: No competing interests.

AUTHORS' CONTRIBUTIONS

PH – was the chief investigator, he conceived the study, led on the protocol development, questionnaire design and writing the manuscript, and approved the final version for publication.

NA, DH, DGT, MGL, BSB, CRC, EM and LV – contributed to protocol development and to writing the manuscript, and approved the final version for publication.

NA – additionally led on the interview studies.

DH – additionally led on the statistical analysis.

PH, NA, DGT, MGL, and CRC - additionally contributed to data acquisition.

CB was the trial manager, contributed to protocol development, questionnaire and database design, and to writing the manuscript, and approved the final version for publication.

AB, JS and TH – contributed to statistical and economic analyses, as well as to writing the manuscript, and approved the final version for publication.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to the following, who are considered members of the INVESTIGATE studies group:

Principal Investigators at other recruiting sites: Mr Andrew Beeby (Queen Elizabeth Hospital, Gateshead, Mr Richard Sill (Wansbeck Hospital, Northumberland), Mr Paul Ballard (South Tees Hospitals).

Research Associates: Dr Megan Murdoch, (Newcastle upon Tyne) and Mr Altaf Mangera and Dr Nadir Osman (Sheffield).

Colleagues contributing to trial management (at the Newcastle Clinical Trials Unit): Mr Chris Speed (senior trial manager), Ms Shelley O'Rourke (trial secretary), Mrs Ruth Wood (database manager), Ms Julie Doughty (randomization service) and Mrs Pauline Potts (assistant database manager) and the staff of *NData* UK Ltd.

Dr Janet Willars who conducted the qualitative interviews, and Mrs Elizabeth Shaw who assisted with analysis of interview data.

Research and clinical nurses: Mrs Angela Black, Ms Liz Dixon, Mrs Wendy Robson (Newcastle upon Tyne), Ms Janet Rose (Gateshead), Mrs Helen Howlett, Mrs Corrine Farrell (Wansbeck), Mrs Lucy Barlow, Mrs Rebecca Thomas (Swansea), Ms Victoria Fowler, Miss Katie Behan (Leicester), Mrs Susannah Hulton (Sheffield), Mrs Julie Potts, Mrs Colette Anderson, Miss Victoria Phelps (South Tees).

And finally, all patients & clinicians who contributed their time to take part in the trial, surveys & interview studies included in INVESTIGATE-I.

FUNDING

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme (reference 09/22/136). The views and

opinions expressed herein are those of the authors and do not necessarily reflect those of the Department of Health. The authors also acknowledge the support of the NIHR, through the Comprehensive Clinical Research Network.

REFERENCES

1. Farage MA, Miller KW, Berardesca E, Maibach HI. Psychosocial and societal burden of incontinence in the aged population: a review. Arch Gynecol Obstet. 2008;277(4):285-90.

2. Hajjar RR. Psychosocial impact of urinary incontinence in the elderly population. Clin Geriat Med. 2004;20(3):553-64.

3. Shaw C. A review of the psychosocial predictors of help-seeking behaviour and impact on quality of life in people with urinary incontinence. J Clin Nurs. 2001;10(1):15-24.

4. Wyman JF, Harkins SW, Choi SC, Taylor JR, Fantl JA. Psychosocial impact of urinary incontinence in women. Obstet Gynecol. 1987;70(3.1):378-81.

5. Moore KH, Wagner TH, Subak L, de Wachter S, Dudding T, Hu TW. Economics of urinary & faecal incontinence and prolapse. In: Abrams P, Khoury S, Cardozo L, Wein A, editors. Incontinence - 5th International Consultation on Incontinence, Paris, 2012. Arnhem, Netherlands: ICUD-EAU; 2013. p. 1832-62.

6. Milsom I, Altman D, Cartright R, Lapitan M, Nelson R, Sillen U, *et al.* Epidemiology of urinary incontinence (UI) and other lower urinary tract symptoms (LUTS), pelvic organ prolapse (POP) and anal incontinence (AI). In: Abrams P, Khoury S, Cardozo L, Wein A, editors. Incontinence - 5th International Consultation on Incontinence, Paris, 2012. Arnhem, Netherlands: ICUD-EAU; 2013. p. 15-108.

7. Hannestad YS, Rortveit G, Sandvik H, Hunskaar S. A community-based epidemiological survey of female urinary incontinence: the Norwegian EPINCONT study. Epidemiology of Incontinence in the County of Nord-Trondelag. Journal of Clinical Epidemiology. 2000;53(11):1150-7.

8. Adekanmi OA, Edwards GJ, Barrington JW. The variation in urodynamic practice in the United Kingdom. J Obstet Gynecol. 2002;22(1):48-50.

9. National Collaborating Centre for Women's & Children's Health. Urinary Incontinence - the management of urinary incontinence in women. London: Commissioned by the National Institute for Health & Clinical Excellence 2006.

10. National Collaborating Centre for Women's & Children's Health. Urinary Incontinence in women - update. London: Commissioned by the National Institute for Health & Care Excellence 2013.

11. National Institute for Health & Care Excellence. Urinary Incontinence - the management of urinary incontinence in women. London: National Institute for Health & Care Excellence 2013.

12. National Institute for Health & Clinical Excellence. Urinary Incontinence - the management of urinary incontinence in women. London: National Institute for Health and Clinical Excellence 2006.

13. Hilton P. Long-term follow up studies in pelvic floor dysfunction: the Holy Grail or a realistic aim? BJOG. 2008;115(2):135-43.

14. Hospital Episode Statistics [database on the Internet]. Department of Health. 2014 [accessed 26th April 2014]. Available from: http://www.hesonline.nhs.uk.

15. Department of Health. Payment by Results in the NHS: tariff for 2012 to 2013. 2013 [updated 25/3/2013]; Available from: https://www.gov.uk/government/publications/confirmation-of-payment-by-results-pbr-arrangements-for-2012-13

16. Medicines and Healthcare Products Regulatory Agency. A summary of the evidence on the benefits and risks of vaginal mesh implants. London: Medicines and Healthcare Products Regulatory Agency, 2014.

17. van Leijsen SA, Kluivers KB, Mol BW, Broekhuis SR, Milani FL, Bongers MY, *et al.* Can Preoperative Urodynamic Investigation be Omitted in Women With Stress Urinary Incontinence? A Non-Inferiority Randomized Controlled Trial. Neurourol Urodyn. 2012;31(7):1118–23.

18. van Leijsen SA, Kluivers KB, Mol BW, Broekhuis SR, Milani FL, van der Vaart CH, *et al.* Protocol for the value of urodynamics prior to stress incontinence surgery (VUSIS) study: a multicenter randomized controlled trial to assess the cost effectiveness of urodynamics in women with symptoms of stress urinary incontinence in whom surgical treatment is considered. BMC Women's Health. 2009;9:22. http://www.controlled-trials.com/mrct/trial/474127/vierhout.

19. Nager CW, Brubaker L, Litman HJ, Zyczynski HM, Varner RE, Amundsen C, *et al.* A randomized trial of urodynamic testing before stress-incontinence surgery. N Engl J Med. 2012;366(21):1987-97.

20. Cohen J. Statistical Power Analysis for the Behavioural Sciences. 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.

21. van Leijsen SA, Kluivers KB, Mol BW, Hout J, Milani AL, Roovers JP, *et al.* Value of urodynamics before stress urinary incontinence surgery: a randomized controlled trial. Obstet Gynecol. 2013;121(5):999-1008.

22. Giarenis I, Cardozo LD. What is the value of urodynamic studies before stress incontinence surgery? BJOG. 2013;120(2):130-2.

23. Brubaker L. An evidence-based approach to urodynamic testing. BJOG. 2013;120(2):127-9.

24. Clement KD, Lapitan MCM, Omar MI, Glazener CMA. Urodynamic studies for management of urinary incontinence in children and adults. Cochrane Database of Systematic Reviews. 2013(10):Art. No.: CD003195.

25. Martin JL, Williams KS, Abrams KR, Turner DA, Sutton AJ, Chapple C, *et al.* Systematic review and evaluation of methods of assessing urinary incontinence. Health Technol Assess. 2006;10(6):1-132, iii-iv.

26. Griffiths D, Kondo A, Bauer S, Diamant N, Liao L, Schafer W, *et al.* Dynamic testing. In: Abrams P, Cardozo L, Khoury S, Wein A, editors. Incontinence - 3rd International Consultation on Incontinence. Plymouth: Health Publications Ltd; 2005. p. 585-674.

27. Rosier PFWM, Kuo H-C, de Gennaro M, Kakizaki H, Hashim H, van Meel TD, *et al.* Urodynamic testing. In: Abrams P, Khoury S, Cardozo L, Wein A, editors. Incontinence - 5th International Consultation on Incontinence, Paris, 2012. Arnhem, Netherlands: ICUD-EAU; 2013. p. 429-506.

28. Buckley BS, Grant AM, Tincello DG, Wagg AS, Firkins L. Prioritising reseach: patients, carers and clinicians working together to identify and prioritise clinical uncertainties in urinary incontinence. Neurourol Urodyn. 2010;29(5):708-14.

29. James Lind Alliance. Urinary incontinence - tackling treatment uncertainties together. Report of the final priority setting workshop 6 November 2008 2009. Available from: http://www.lindalliance.org/pdfs/JLA UI Report 18 3 08_FINAL_amended.pdf.

30. Weber AM, Taylor RJ, Wei JT, Lemack G, Piedmonte MR, Walters MD. The cost-effectiveness of preoperative testing (basic office assessment vs. urodynamics) for stress urinary incontinence in women. BJU Int. 2002;89(4):356-63.

31. Brookes ST, Donovan JL, Wright M, Jackson S, Abrams P. A scored form of the Bristol Female Lower Urinary Tract Symptoms questionnaire: data from a randomized controlled trial of surgery for women with stress incontinence. Am J Obstet Gynecol. 2004;191(1):73-82.

32. Ward K, Hilton P. Multicentre randomised trial of TVT and colposuspension: 5 year follow-up. BJOG. 2008;115(2):226-33.

33. MacLean AB, Cardozo L. Recommendations arising from the 42nd Study Group. In: MacLean AB, Cardozo L, editors. Incontinence in Women. London: RCOG Press; 2002. p. 433-41.

34. RCOG. Surgical treatment of urodynamic stress incontinence. Guideline No 35 2003. Available from: http://www.rcog.org.uk/files/rcog-corp/uploaded-files/GT35UrodynamicStressIncontinence2003.pdf.

35. Fowler G, Richmond D. Urodynamics: a mandatory preoperative investigation? The Obstetrician and Gynaecologist. 2006;8(2):86-90.

36. Basu M, Duckett JR, Moran P, Freeman R. Clinicians' views on the NICE guideline on the management of female urinary incontinence. J Obstet Gynaecol. 2009;29(6):529-32.

37. Agur W, Housami F, Drake M, Abrams P. Could the National Institute for Health and Clinical Excellence guidelines on urodynamics in urinary incontinence put some women at risk of a bad outcome from stress incontinence surgery? BJU Int. 2009;103(5):635-9.

38. Majumdar A, Latthe P, Toozs-Hobson P. Urodynamics prior to treatment as an intervention: a pilot study. Neurourol Urodyn. 2010;29(4):522-6.

39. Murdoch M, McColl E, Howel D, Deverill M, Buckley BS, Lucas MG, *et al.* INVESTIGATE-I (INVasive Evaluation before Surgical Treatment of Incontinence Gives Added Therapeutic Effect?): study protocol for a mixed methods study to assess the feasibility of a future randomised controlled trial of the clinical utility of

invasive urodynamic testing. Trials. 2011; 12:169 [Open access]. Available from: http://www.trialsjournal.com/content/12/1/169.

40. Armstrong N, Hilton P. Doing diagnosis: whether and how clinicians use a diagnostic tool of uncertain clinical utility. Soc Sci Med. 2014;120(2014):208-14.

41. Hilton P, Bryant A, Howel D, McColl E, Buckley BS, Lucas MG, *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(8):1223-30.

42. Homer T, Vale L, Shen J, Hilton P, on behalf og the INVESTIGATE studies group. The results of a cost-utility analysis and a value of information analysis on the feasibility of a future randomised control trial on invasive urodynamic testing prior to surgery for stress urinary incontinence in women. Proceedings of International Society For Pharmacoeconomics and Outcomes Research (ISPOR), 17th Annual European Congress; 2014 November 2014; Amsterdam, Netherlands.

43. Department of Health. NHS Research Governance Framework for Health and Social Care 2005. Available from: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4122427.pdf.

44. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol. 2010;10:67.

45. Bugge C, Williams B, Hagen S, Logan J, Glazener C, Pringle S, *et al.* A process for Decision-making after Pilot and feasibility Trials (ADePT): development following a feasibility study of a complex intervention for pelvic organ prolapse. Trials. 2013;14:353.

46. Charlesworth G, Burnell K, Hoe J, Orrell M, Russell I. Acceptance checklist for clinical effectiveness pilot trials: a systematic approach. BMC Med Res Methodol. 2013;13:78.

47. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract. 2004;10(2):307-12.

48. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, *et al.* A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol. 2010;10:1.

49. Ramsay IN, Ali HM, Hunter M, Stark D, Donaldson K. A randomized controlled trial of urodynamic investigations prior to conservative treatment of urinary incontinence in the female. Int Urogynecol J & Pelvic Floor Dysfunct. 1995;6(5):277-81.

50. Allahdin S, Bain C, Glazener C. Feasibility study for a randomised controlled trial evaluating the use of absorbable mesh, polydioxanone and polyglactin sutures for anterior and posterior vaginal wall prolapse repairs. International Continence Society; Christchurch, New Zealand 2006. p. 412 (abstract).

51. Kitchener HC, Dunn G, Lawton V, Reid F, Nelson L, Smith AR. Laparoscopic versus open colposuspension--results of a prospective randomised controlled trial. BJOG. 2006;113(9):1007-13.

52. Hagen S, Stark D, Glazener C, Sinclair L, Ramsay I. A randomized controlled trial of pelvic floor muscle training for stages I and II pelvic organ prolapse. Int Urogynecol J & Pelvic Floor Dysfunct. 2009;20(1):45-51.

53. Jenkinson C, Layte R. Development and testing of the UK SF-12 (short form health survey). J Health Serv Res Policy. 1997;2(1):14-8.

54. Brooks R. EuroQol: the current state of play. Health Policy. 1996;37(1):53-72.

55. Avery K, Donovan J, Peters TJ, Shaw C, Gotoh M, Abrams P. ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. Neurourol Urodyn. 2004;23(4):322-30.

56. Kelleher CJ, Cardozo LD, Khullar V, Salvatore S. A new questionnaire to assess the quality of life of urinary incontinent women. BJOG. 1997;104(12):1374-9.

57. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fanti JA. Health-related quality of life measures for women with urinary incontinence: The incontinence impact questionnaire and the urogenital distress inventory. Quality of Life Research. 1994;3(5):291-306.

58. Glaser BG, Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago: Aldine; 1967.

59. Hilton P, McColl E, Howel D, Deverill M, Buckley BS, Lucas MG, *et al.* INVESTIGATE-I (INVasive Evaluation before Surgical Treatment of Incontinence Gives Added Therapeutic Effect?): a pragmatic multicentre pilot study to assess the feasibility of a future randomised controlled trial. Protocol version 1.2 2012. Available from: http://www.nets.nihr.ac.uk/__data/assets/pdf_file/0017/54350/PRO-09-22-136.pdf.

60. Shanyinde M, Pickering RM, Weatherall M. Questions asked and answered in pilot and feasibility randomized controlled trials. BMC Med Res Methodol. 2011;11:117.

61. McCann SK, Campbell MK, Entwistle VA. Reasons for participating in randomised controlled trials: conditional altruism and considerations for self. Trials. 2010;11:31.

62. Rosenbaum JR, Wells CK, Viscoli CM, Brass LM, Kernan WN, Horwitz RI. Altruism as a reason for participation in clinical trials was independently associated with adherence. J Clin Epidemiol. 2005;58(11):1109-14.

63. Truong TH, Weeks JC, Cook EF, Joffe S. Altruism among participants in cancer clinical trials. Clin Trials. 2011;8(5):616-23.

64. Fletcher B, Gheorghe A, Moore D, Wilson S, Damery S. Improving the recruitment activity of clinicians in randomised controlled trials: a systematic review. BMJ Open. 2012;2(1):e000496.

65. Rolstad S, Adler J, Ryden A. Response burden and questionnaire length: is shorter better? A review and meta-analysis. Value Health. 2011;14(8):1101-8.

66. Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I, *et al.* Methods to increase response to postal and electronic questionnaires. Cochrane Database of Systematic Reviews. 2009(3):Art. No.: MR000008.

67. Brueton VC, Tierney JF, Stenning S, Meredith S, Harding S, Nazareth I, *et al.* Strategies to improve retention in randomised trials: a Cochrane systematic review and meta-analysis. BMJ Open. 2014;4(2):e003821.

68. Cook JA, Hislop J, Adewuyi TE, Harrild K, Altman DG, Ramsay C, *et al.* Assessing methods to specify the target difference for a randomised controlled trial – DELTA (Difference ELicitation in TriAls) review. Health Technol Assess. 2014;18(28).

Table 1 – Summary of findings against 14 methodological issues for feasibility research

Submitted as separate file

Table 2: Screening & recruitment numbers

including screening codes (1-15) for those women not randomized, sorted by overall frequency of reporting of codes

Code	Description	Total	Per cent
11	Patient has not undergone a course of pelvic floor training	105	14%
14	Urge incontinence	92	12%
13	Other (give details)	86	11%
15	Patient did not attend clinic	81	11%
7	Patient does not wish to participate, include reason if offered	59	8%
1	Symptomatic utero-vaginal prolapse requiring treatment	40	5%
8	Clinician feels surgery inappropriate	39	5%
9	Patient does not wish surgery	21	3%
2	Previous surgery for urinary incontinence or pelvic organ prolapse	9	1%
3	Urodynamic investigation within the last three days	7	1%
10	Patient does not consider her family is complete	6	1%
4	Neurological disease causing urinary incontinence	1	0%
5	Current involvement in a conflicting research study	0	0%
6	Unable to give competent informed consent	0	0%
12	Study not discussed at clinic visit (please give reason)	3	0%
	Recruited	**222**	**29%**
	Total screened	**771**	100%
	Screened women recruited	222/771	29%
	Eligible women recruited	222/284	78%

Table 3: Summary of demographic data at baseline by trial arm

	'IUT'		'no IUT'	
	n	%	n	%
Ethnicity				
Caucasian	110	99%	106	97%
Black	0	0%	0	0%
Asian	1	1%	3	3%
Other	0	0%	0	0%

	'IUT'				'no IUT'			
	n	mean (SD)	median (IQR)	range	n	mean (SD)	median (IQR)	range
Age	112	47.1(9.5)	46.5 (40-52)	29-75	110	46.8 (10.0)	46.5 (40-52)	24-77
BMI	106	29.3 (6.5)	28.3 (24.4-33.7)	20-55	102	27.4 (5.0)	26.8 (23.9-30.7)	18-45

BMI=body mass index; SD=standard deviation; IQR=interquartile range; 'IUT'=invasive urodynamic testing (intervention) arm; 'no IUT'=no invasive urodynamic testing (control) arm.

Table 4: Summary of numeric outcome measures by trial arm and data collection time-point

	'IUT'								'no IUT'								Overall completion rate[1]			
	Baseline				6 months				Baseline				6 months				Baseline		6 months[2]	
Questionnaire (*possible scores*)	n	Mean (SD)	Median (IQR)	Range	n	Mean (SD)	Median (IQR)	Range	n	Mean (SD)	Median (IQR)	Range	n	Mean (SD)	Median (IQR)	Range	Partial n (%)	Complete n (%)	Partial n (%)	Complete n (%)
ICIQ-FLUTS Overall score (*0-48*)	77	16.9 (5.7)	17 (13-21)	4-37	47	9.2 (7.5)	8 (4-12)	0-38	85	16.4 (6.3)	16 (11-21)	3-34	66	6.9 (5.0)	6 (3-9)	0-26	3 (2)	162 (98)	5 (4)	113 (90)
Subscales:																				
Filling (*0-16*)	78	4.4 (2.3)	4 (3-6)	0-11	48	3.0 (2.3)	3 (1-4)	0-11	85	4.0 (2.6)	3 (2-6)	0-10	66	2.4 (1.8)	2 (1-3)	0-8	2 (1)	163 (99)	3 (3)	114 (91)
Voiding (*0-12*)	79	1.8 (2.0)	1 (0-3)	0-9	49	2.0 (2.0)	2 (0-3)	0-9	86	1.5 (1.7)	1 (0-2)	0-9	68	2.3 (2.1)	2 (0-4)	0-8	0 (0)	165 (100)	1 (1)	117 (94)
Incontinence (*0-20*)	78	10.8 (3.3)	11 (8-13)	2-19	49	4.0 (4.9)	3 (1-5)	0-20	86	10.8 (3.6)	11 (8-13)	2-19	68	2.3 (3.1)	2 (0-3)	0-16	1 (1)	164 (99)	1 (1)	117 (94)
ICIQ-UI SF (*0-21*)	78	14.0 (3.7)	14 (12-16)	4-21	49	5.3 (6.0)	3 (0-8)	0-21	85	14.1 (3.8)	15 (12-17)	4-21	65	3.3 (4.5)	1 (0-4)	0-18	2 (1)	163 (99)	3 (3)	114 (91)
ICIQ-LUTSqol (*19-76*)	73	46.8 (10.9)	47 (40-52)	26-74	44	26.7 (12.3)	22 (20-28)	19-76	84	48.5 (11.7)	46 (39-58)	30-72	65	25.3 (9.6)	21 (20-28)	19-65	8 (5)	157 (95)	9 (7)	109 (87)
UDI Overall score (*0-300*)	64	133.3 (43.5)	133.5 (109-159)	25-245	42	49.1 (44.1)	37.1 (17-69)	0-191	74	130.1 (43.8)	125.8 (96-162)	50-227	59	33.9 (39.7)	24.2 (4-46)	0-150	27 (16)	138 (84)	17 (14)	101 (81)
Subscales:																				
Stress (*0-100*)	76	82.9 (21.0)	87.5 (75-100)	25-100	50	24.5 (26.1)	25 (0-38)	0-100	80	80.2 (21.2)	87.5 (63-100)	38-100	65	18.1 (27.0)	0 (0-25)	0-100	6 (4)	156 (95)	2 (2)	115 (92)
Irritative (*0-100*)	71	38.4 (25.4)	33.3 (17-54)	0-100	48	16.5 (20.5)	8.3 (0-25)	0-100	80	33.7 (24.3)	31.3 (17-50)	0-92	64	10.0 (13.3)	4.2 (0-17)	0-54	13 (8)	151 (91)	6 (5)	112 (90)
Obstructive/ discomfort (*0-100*)	68	17.6 (17.6)	13.6 (6-23)	0-73	43	10.9 (15.1)	4.6 (0-18)	0-64	80	14.8 (14.2)	13.6 (3-20)	0-61	64	8.9 (12.4)	2.3 (0-14)	0-57	17 (10)	148 (90)	11 (9)	107 (86)

[1] Complete responses are defined as women who completed all questions on the particular questionnaire scale, and partial responses as those who completed at least one question but did not fully complete the particular scale.

[2] In addition to complete and partial responses, there were seven completely blank questionnaires amongst the six-month responses.

ICIQ-FLUTS=International Consultation on Incontinence Female Lower Urinary Tract Symptoms questionnaire; UDI=Urogenital Distress Inventory; SD=standard deviation; IQR=interquartile range; 'IUT'=invasive urodynamic testing (intervention) arm; 'no IUT'=no invasive urodynamic testing (control) arm.

Table 5: Summary statistics for paired changes in scale scores (baseline - six-month)

Questionnaire	n	Mean (SD)	Median (IQR)	Range
'IUT' arm				
ICIQ-FLUTS -Overall score	31	7.8 (5.9)	7 (4-15)	-5 to +18
ICIQ-UI SF	34	8.9 (6.0)	11 (4-13)	-3 to +16
ICIQ-LUTSqol	29	20.0 (11.4)	23 (12-28)	-5 to +41
UDI - Overall score	27	79.5 (45.5)	75 (51-122)	-21 to +161
'no IUT' arm				
ICIQ-FLUTS -Overall score	48	9.3 (7.3)	10.5 (5.5 – 15)	-9 to +22
ICIQ-UI SF	49	10.2 (5.8)	11 (6-15)	-4 to +21
ICIQ-LUTSqol	47	23.7 (13.9)	23 (14-35)	-3 to +50
UDI - Overall score	41	94.1 (55.3)	92 (70 – 117)	-66 to +221

ICIQ-FLUTS=International Consultation on Incontinence modular questionnaires: Female Lower Urinary Tract Symptoms questionnaire; ICIQ-UI SF=ICIQ Urinary Incontinence Short Form questionnaire; ICIQ-LUTSqol=ICIQ Lower Urinary Tract Symptoms quality of life questionnaire; UDI=Urogenital Distress Inventory; SD=standard deviation; IQR=interquartile range; 'IUT'=invasive urodynamic testing (intervention) arm; 'no IUT'=no invasive urodynamic testing (control) arm.

Table 6: Total numbers necessary in definitive trial when analysis compares mean changes in ICIQ-FLUTS total score over six months

	Difference to be detected		
	2	**3**	**4**
Number of RESPONSES to primary outcome	516	230	130
Number of RECRUITED patients	922	410	232
Number of eligible women APPROACHED	1182	526	298
Number of women SCREENED for eligibility	3194	1422	806

ICIQ-FLUTS=International Consultation on Incontinence modular questionnaires: Female Lower Urinary Tract Symptoms questionnaire

Figure 1: Trial CONSORT flow diagram

'IUT'=invasive urodynamic testing (intervention) arm; 'no IUT'=no invasive urodynamic testing (control) arm; DNA=did not attend.

Submitted as separate file

Figure 2: Monthly target and actual recruitment numbers

The original and revised predictions of overall recruitment are shown as continuous and dashed lines, and actual recruitment in histogram; the overall Comprehensive Local Research Network (CLRN) black/red/amber/green flag or 'recruitment to target' status is also illustrated.

Submitted as separate file

Figure 1: Trial CONSORT flow diagram

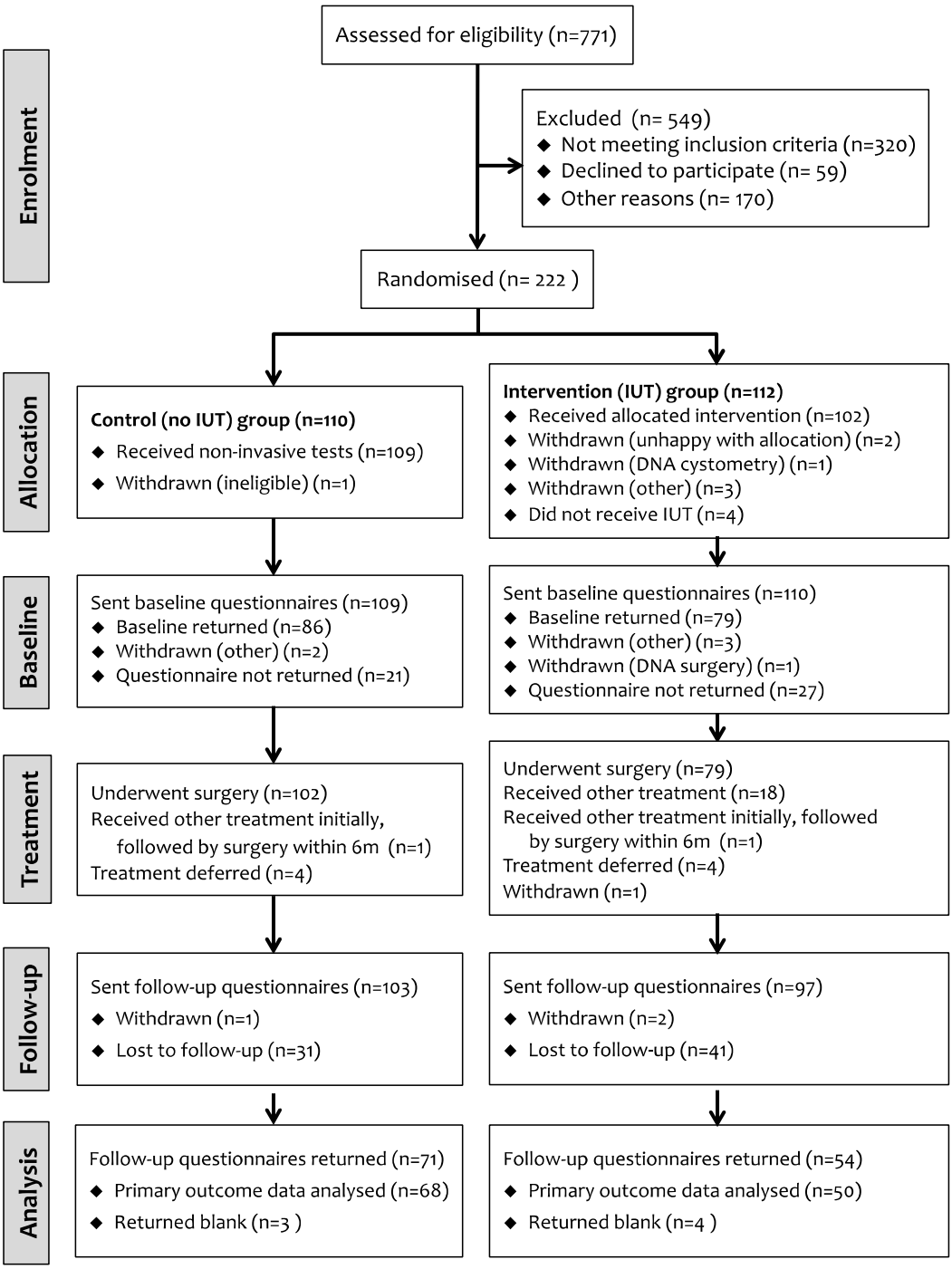
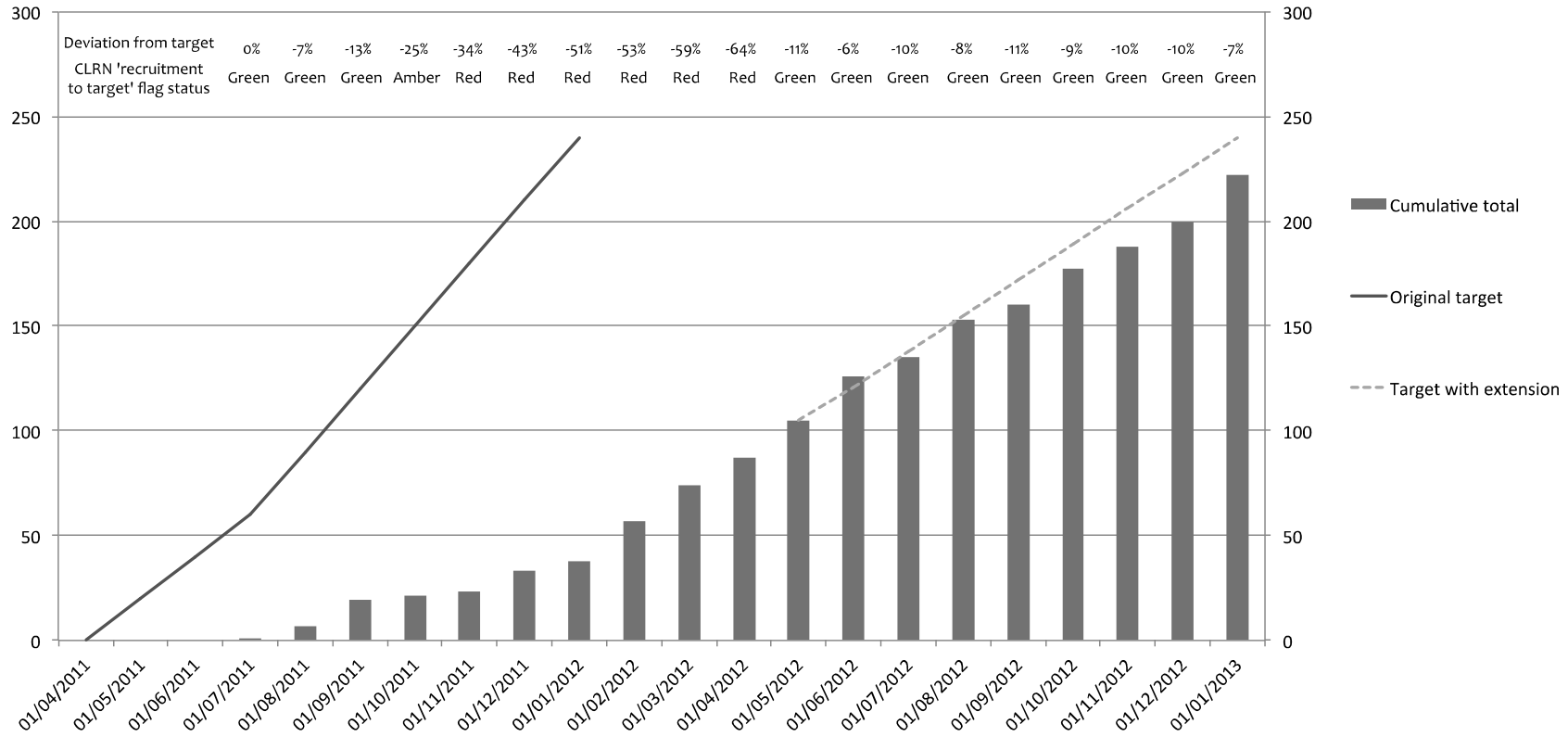


Figure 2

Figure 2: Monthly target and actual recruitment numbers

The original and revised predictions of overall recruitment are shown as continuous and dashed lines, and actual recruitment in histogram; the overall Comprehensive Local Research Network (CLRN) black/red/amber/green flag or 'recruitment to target' status is also illustrated.



Additional files provided with this submission:

Additional file 1: Table 1.1.docx, 97K
http://www.trialsjournal.com/imedia/1841363040181562/supp1.docx
Additional file 2: CONSORT Checklist 020715.doc, 217K
http://www.trialsjournal.com/imedia/8296303691853479/supp2.doc