CONFIDeNT

Statistical Analysis Plan

Version: 1.0
Date: 19/Jun/2014

Person(s) contributing to the analysis plan	
Name(s) and position(s)	Dr Stephen Bremner, trial statistician Prof Sandra Eldridge, director of PCTU Prof Charles Knowles, chief investigator Emma Horrocks, trial coordinator
Authorisation	
Position	**Chief or principal investigator**
Name	Prof Charles Knowles
Signature	
Date	DD/MMM/YYYY
Position	**Senior trial statistician**
Name	Prof. Sandra Eldridge
Signature	
Date	DD/MMM/YYYY

1. INTRODUCTION

Purpose of statistical analysis plan

The purpose of this document is to provide details of the statistical analyses and presentation of results to be reported within the principal paper(s) of the CONFIDeNT trial. Subsequent papers of a more exploratory nature (including those involving baseline data only) will not be bound by this strategy but will be expected to follow the broad principles laid down in it. Any exploratory, post hoc or unplanned analyses will be clearly identified in the respective study analysis report.
The structure and content of this document provides sufficient detail to meet the requirements identified by the International Conference on Harmonisation (ICH) and the PCTU SOP (PCTU/07).
The following were reviewed in preparation for writing this document:
Trial application submitted 04/01/2011
ICH E9 Guidance on statistical principals for clinical trials
ICH E3 Structure and content of clinical study reports
CONSORT guidelines for the reporting of randomised trials

Members of the writing committee

Stephen Bremner (SB) and Sandra Eldridge (SE) were primarily responsible for writing the Statistical Analysis Strategy with SB responsible for writing the computer code implementing the analysis strategy and implementing the strategy at the point of analysis. Emma Horrocks and Prof Charles Knowles helped refine outcome definitions and choose variables for the multiple imputation. This document was developed prior to examination of unblinded trial data and will not be implemented prior to final approval.

Summary

DESIGN: Pragmatic multi-centre, double-blinded, placebo-controlled trial of 227 patients randomised to receive the intervention (PTNS) or sham (needle insertion and electrical stimulation). All patients follow an assessment period, recruitment, allocation, standard 3 month treatment protocol (one 30. min session per week for 12 weeks) with trial outcomes determined at 14 weeks.
SETTING: 19 UK centres providing specialist nurse-led treatment for pelvic floor disorders.
TARGET POPULATION: Patients aged > 18 years with faecal incontinence (FI) who have failed conservative treatments and whose symptoms are sufficiently severe to merit further intervention (80-90% female based on departmental data).
HEALTH TECHNOLOGIES BEING ASSESSED: PTNS (Urgent ® PC neuromodulation system) is produced by a single manufacturer (Uroplasty ®). The equipment includes a hand held pulse generator unit, single use leads and fine needle electrodes. Needle insertion is performed in a sitting position in an outpatient setting on either leg adhering to the manufacturer's protocol (and specialist training). Treatment is for 30 mins. weekly for a duration of 12 weeks. Validated sham stimulation - insertion of the Urgent PC needle subcutaneously at the same site with electrical stimulation delivered to the distal foot using TENS.

MEASUREMENT OF COSTS AND OUTCOMES: Primary outcome variable: change in weekly FI episodes (calculated from bowel diaries) expressed as proportion of patients achieving ≥ 50% reduction in FI episodes per week; Secondary outcomes: (1) percentage change in FI episodes per week, (2) change in mean number of FI episodes per week, (3) validated patient-rated quantitative outcomes including symptom severity score (St Mark's score), disease-specific: FI-QOL, and generic: EQ-5D QOL measures, and SF-36, (4) FI-specific patient-centred outcomes (5 validated key issues), (5) Likert scales of patient's global impression of success (0-10), (6) Short urinary symptom assessment. Adverse events and anti-diarrhoeal drug usage will also be recorded. Economic analysis will measure direct NHS costs with utilities derived from the EQ-5D. The proposed HE analysis will be detailed in a separate document.

Changes from planned analysis in the protocol

We decided to fit random centre effects rather than fixed effects on the basis of findings by Kahan & Morris (Kahan & Morris, 2013). We also decided to multiply impute the data and remove any reference to last observation carried forward.

STUDY OBJECTIVES AND ENDPOINTS

Study objectives

Primary objectives

To determine the effectiveness of PTNS versus sham electrical stimulation based on changes in the number of weekly FI episodes from baseline (bowel diary completed over two week period prior to first intervention) to end of treatment (bowel diary completed for weeks 12 and 13)

Secondary objectives

To determine the effectiveness of PTNS versus sham electrical stimulation (TENS) based on changes in validated incontinence scores, patient-centred FI-related symptoms and disease-specific and generic quality of life measures from baseline (bowel diary completed over two week period prior to first intervention) to end of treatment (bowel diary completed for weeks 12 and 13)

Exploratory objectives

None

Outcome measures

Primary outcome

Change in weekly FI episodes expressed as proportion of patients achieving ≥ 50% reduction in FI episodes per week. The change is measured between pre- and post-treatment bowel diaries.

The number of FI episodes per day are the sum of episodes in Q2a (rush) and Q2b (passive leakage) of the bowel diary. The average number per week is the sum of all 14 days, divided by 2.

%change = 100% × (#FI(baseline) - #FI(end of intervention)) / #FI(baseline)

Where #FI is the average number of episodes of FI per week. Where %change is negative, this represents an increase in FI episodes; where it is positive, this represents a decrease in FI episodes.

A patient achieving a ≥ 50% reduction will be classed as a treatment success, otherwise the patient is classed as a treatment failure.

Secondary outcomes

Percentage change in FI episodes per week, from baseline (the two-week period just before the 1st treatment) to end of treatment i.e. bowel diary for two weeks after the 12th treatment; for three additional cut offs: an improvement ≥ 25% vs. less, an improvement ≥ 75% vs. less and an improvement of 100% vs. less

Continuous change in FI episodes per week; i.e. average number of FI episodes per week for the two weeks post-treatment, minus average number of FI episodes per week at baseline (from pre-treatment two-week bowel diary).

Continuous change in FI episodes per week (rush and passive leakage as two separate outcomes); i.e. average number of episodes per week for the two weeks post-treatment, minus average number of episodes per week at baseline (from pre-treatment two-week bowel diary).

St Mark's incontinence score (Vaizey et al. 1999)

Likert scale of patient's global impression of success

Patient-centred FI-related symptoms

Disease specific and generic quality of life measures

- EQ-5D

- SF-36 (8 domains)

- Faecal incontinence quality of life score- four domains: coping, embarrassment, lifestyle and depression

- Gastro-intestinal quality of life score

Short urinary symptom assessment (descriptive)

Change in medication use: has pad usage/loperamide usage decreased/remained the same/increased? (descriptive)

Safety outcomes

None

STUDY METHODS

Overall study design and plan

Target for analysis: 106 intervention and 106 sham participants
Actual number randomised: 227
Date of first randomisation: 21/01/2012
Date of last randomisation: 31/10/2013
Trial design: individually randomised, parallel group
Blinding: See section 3.4
Randomised Interventions: PTNS vs. sham
Allocation ratio: 1:1

Selection of study population

Inclusion Criteria

Faecal incontinence sufficiently severe enough to warrant intervention

Failure of appropriate conservative therapies

Age ≥ 18 years

Exclusion Criteria

Inability to provide informed consent for the research study

Inability to fill in the detailed bowel diaries required for outcome assessments (this will exclude participants who do not speak / read English)

Neurological diseases, such as diabetic neuropathy, multiple sclerosis and Parkinson's disease (any participant with painful peripheral neuropathy)

Anatomical limitations that would prevent successful placement of needle electrode

Other medical conditions precluding stimulation: e.g. bleeding disorders, certain cardiac pacemakers, peripheral vascular disease or ulcer, lower leg cellulitis

Congenital anorectal anomalies or absence of native rectum due to surgery

A cloacal defect

Present evidence of external full thickness rectal prolapse

Previous rectal surgery (rectopexy / resection) done < 12 months ago (24 months for cancer),

Stoma *in situ*

Chronic bowel diseases such as inflammatory bowel disease leading to chronic uncontrolled diarrhoea

Pregnancy or intention to become pregnant

Previous experience of SNS or PTNS

Method of treatment assignment and randomisation

Participants were randomised, with allocation concealment, at a ratio of 1:1 at visit 2 using a web-based computer programme to receive either PTNS or sham. This was performed by the Nottingham Clinical Trials Unit Study. Centres inputted the sex of the participant. Sex was used to reduce the potential confounding effects of variation in outcomes between male and female participants. Males represent approximately 10% of patients and have differing pelvic physiology and often disease aetiology (e.g. post anal surgery rather than childbirth). As only 1 or 2 male patients were expected to be enrolled from each centre, randomisation was first stratified on sex, and then within females only, further stratified on centre reducing the possibility that all the males are allocated to PTNS by chance. Randomly permuted blocks of length randomly varying 2, 4 and 6 will be used to ensure near balance between PTNS and sham arms.

Treatment masking (Blinding)

Blinding of patients: For both interventions: (1) a standardised description of the techniques were read from a card. This described an electrical sensation variably in the ankle or foot with or without motor responses in the foot (note: there is significant variability in conscious sensation and motor responses even between patients undergoing only PTNS); (2) the equipment (identical for both interventions) was shown to the patient; (3) the lower extremity was be draped from view; and (4) the audible sounds produced by the Urgent PC unit identical.
Performance bias considerations: Since the sham group might be expected to seek more advice than the treatment arm (if the hypothesis that PTNS is more effective than placebo is correct), the interaction of the administering nurse/physician was standardised so that general supportive advice given at consultations was identical for all participants. This was limited to a general welcome, answers to any concerns (whilst recording adverse events), advice on loperamide dosages and pad use (both recorded in outcome variables).
Blinding of trial staff: two members of staff were available at each site to run the study. Randomisation into the treatment or placebo arm of the study occurred at Visit 2, after all the documentation had been completed. At this point, the member of staff carrying out the PTNS or sham was unblinded. That same staff member carried out all 12 treatments for the patient. Following the final treatment the member of staff who remained blinded collected all of the final data, before allowing the patient to find out if they were in the sham or treatment arm. In this way, the staff member conducting the final meeting with the patient remained blinded until the end.

Sample size determination

Research into treatment of FI is currently hampered by the lack of a valid and reliable tool that allows standardisation of outcomes. There are advantages and disadvantages of the numerous possible quantitative outcome variables e.g. individual symptoms and composite scores, and generally poor correlation of either with disease specific or generic quality of life measures. Of possible outcomes, the most frequently used and

probably least affected by subjective reporting differences is number of FI episodes per unit time (usually per week). This outcome, obtained directly from the mean of 2 week bowel diary frequencies has been employed in almost all contemporary studies of FI interventions including recent SNS studies. The problem with this variable is that, being a count, it has a Poisson distribution and is over-dispersed i.e. has greater variability than expected. This raises major difficulties in defining a *clinically significant* mean reduction in FI episodes per week in a population of patients with widely dispersed starting FI frequencies. To counter this problem, almost all contemporary studies have adopted a primary outcome using a categorical measure of percentage reductions i.e. the proportion of patients who have a 50% or greater reduction in faecal incontinence episodes per week. We justify this approach on the following basis:

The most important inferred outcome of this study will be the comparison of PTNS outcomes with that of other interventional treatments especially those of SNS. Since the primary outcome of nearly all studies of SNS has been based on the $\geq$ 50% reduction in FI episodes rule, the continued use of this outcome will better inform bodies such as NICE. Indeed, this outcome was used in the NICE ruling on sacral nerve stimulation; it was also the primary outcome in the 16-site multicentre FDA investigational device exemption (IDE) trial of sacral nerve stimulation in 120 patients with FI.

This outcome has also been the approach of choice for urinary incontinence episodes in the only pivotal trial of PTNS in the urology literature and also for NICE commissioned systematic reviews.

Baseline and post treatment FI episodes expressed as continuous variables yield data from over-dispersed Poisson distributions. The arithmetic means of these variables are very difficult to correlate with significant clinical effect e.g. a mean change of 5 FI episodes per week is not possible in patients with starting frequencies of four or fewer and is of little or no benefit to a patient with a starting frequency of 50. The change variable will however remain a secondary outcome.

Previous publications and our own data on 50 patients suggest a 60% success rate for PTNS on the basis of above justified primary outcome measure. There are no RCT data for PTNS in FI. However the pivotal level I SUmiT trial of PTNS in overactive bladder symptoms (OAB) which used a similar global response assessment of urinary incontinence and an intention to treat analysis, observed a moderate or marked improvement in symptoms in 55% PTNS group and only 21% sham group. On the basis that placebo responses are frequently higher for bowel rather than bladder symptoms we have selected a sham response rate of 35% whilst keeping this more conservative estimate of treatment response of 55%. We believe this difference remains clinically important. Two hundred and twelve patients are required to detect this difference with 80% power at the 5% significance level. We aimed recruit 235 patients at baseline to allow for a 10% failure to attend the 2nd visit (allocation and first intervention).

DATA COLLECTION

Baseline

Age, sex, history of faecal incontinence (including type), urinary symptom history, previous faecal incontinence treatments, medications, past medical history, past obstetric history. (see CRF 2 v5 for items)
Bowel diary (14 consecutive days)
Gastrointestinal quality of life index, patient centred FI symptoms, SF-36 Health Survey, QoL scale for faecal incontinence, St. Mark's faecal incontinence score, EQ-5D health questionnaire (see CRF 3 v3 for detail)

Follow up

Visits 2-13: PTNS or tens machine settings and response (sensory or motor), adverse events, any change in pad usage, any changes in medication use? (see CRF 5 v3)
Week 7: bowel diary over 7 consecutive days
Final visit (post treatment) (week 14): which treatment did patient think they were on? Any effect on urinary symptoms? Any change in loperamide or codeine use? Any change in pad use? How patient felt before, during and after treatment, Likert scale of success.
Bowel diary over 14 consecutive days
CRF 3 v3

Timing of data collection

Event	Visit 1	Telephone Conversation	Visit 2	Visits 3-13	Visit 14
Eligibility assessment	X				
Bowel Diary		X		Visit 7-8	X
Consent			X		
Participant Contact Information Sheet			X		
Eligibility assessment (CRF1)			X		
Initial assessment (CRF2)			X		
Questionnaires (CRF3)			X		X
Randomisation			X		
Randomisation information (CRF4)			X		
Intervention			X	X	
Record stimulation parameters adverse events and medication / pad usage (CRF5)			X	X	
Adverse Events Log			X	X	X
Concomitant Medications Log			X	X	X
Post treatment Information (CRF6)					X
Final Study Visit Information (CRF7)					X

GENERAL ISSUES FOR STATISTICAL ANALYSIS

All analyses will be conducted two sided and significance interpreted at the 5% level.

Blinding of the statistical analysis

The trial statistician will remain blind to allocation until this analysis plan was signed off

Analysis populations

Intent-to-treat population

The intention-to-treat (ITT) sample is defined for this trial as all participants randomised into the trial, *who received at least their first treatment*, included in the intervention group to which they were randomised.

Available-case population

N/A

Per protocol population

Patients attending at least 10 treatment sessions in 13 weeks will be classed as treatment completers.

Safety population

N/A

Other populations

N/A

Database

Description

Data were entered by the trial manager, clinical academic fellow and data entry clerks onto a Microsoft Access 2010 database held on the Barts Cancer Institute secure server. Data was entered at QMUL.

Data quality

Completeness of data was checked each time a report was generated for data monitoring committee meetings, and prior to final analysis. All eligibility and primary outcome data were checked by a member of the trial team other than the person who entered it. This was done in batches, as and when time permitted. A 10% random sample of CRFs for secondary outcomes were checked and the overall error rate was found to be below the 2% error rate that would have necessitated a 100% check of the secondary outcomes data.

Database freeze and lock

Once the trial team completed all data entry and checking, the data date stamped and frozen for transfer to Stata version 12.1 using the *odbc* facility in Stata. The statistician responsible for the analysis conducted additional data checks. These range

checks, logical and consistency checks which may not have been picked up by checks performed at the individual level.
Discrepancies were dealt with by the trial manager checking the paper CRFs, and the database was locked for analysis i.e. it was transferred to a read only location.

Analysis software

The analysis will be carried out using Stata version 12.1, interfacing with Realcom Impute which will be used to multiply impute missing outcome and baseline covariate data.

Methods for withdrawals, loss to follow-up and missing data

All patients randomised who receive the first treatment will be included in the intention-to-treat analysis of primary endpoint. Prior to the first treatment, it was anticipated that some patients (up to 10%) would fail to attend after eligibility was assessed due to a failure of compliance with the travel and attendance needs of the treatment course or study. These were not counted as study recruits. Those in whom post-treatment data are unavailable at 14 weeks for any reason (loss to follow up, failure to complete treatment) will have their outcome multiply imputed under the assumption of missing at random (MAR) using variables prognostic of outcome, such as measure of outcome made at baseline, and others that are predictive of missingness such as mean number of FI episodes per week at baseline, age, sex, and where available, mid-study bowel diary data. See appendix (b) for details. Multilevel multiple imputation will be performed using the multivariate normal distribution in Realcom Impute, using treatment allocation, patient sex and allocation as auxiliary variables. After a burn in of 1,000 runs of the MCMC sampler, missing values will be filled every 500th run to create a total of 10 completed datasets for analysis. The data will be analysed in Stata and the results pooled by Rubin's rules.

Method for handling centre effects

Study centre will be included as a random effect

Method for handling randomisation stratification or minimisation factors

Patient sex will be included as a fixed effect, study centre as a random effect

Method for handling clustering effects

The intracluster correlation coefficients (ICC) and their 95% confidence intervals for the outcomes by centre (level 2) will be estimated using the user-contributed Stata command *sea_obi* which allows the ICC to be negative. Random effects models will be fitted by restricted maximum likelihood estimation (e.g. *xtmixed ..., reml*). However, should these fail due to the between centre variance being close to zero, random effects models will be fitted by generalised least squares (e.g. *xtregress*). If ICCs are estimated as ≤ 0 then we will use similar regression models without adjusting for clustering.

Method for selecting other variables that will be adjusted for

This was agreed by consensus prior to any data extraction and the decision was: fit fixed effects for sex, randomisation and baseline level of outcome.

Multiple comparisons and multiplicity

No adjustments to p-values planned

Method for handling non-adherence

All patients randomised who receive the first treatment will be included in the intention-to-treat analysis of primary endpoint.

Method for handling time-varying interventions
N/A

Method for handling outliers and influential points
N/A

Data from external sources
N/A

Derived and computed variables

In the bowel diary, participant's record:
Controlled bowel motions: No incontinence – pads or pants remained clean

2. Uncontrolled bowel movements: Incontinence – underwear, pads or pants got dirty.
Within this section patients are asked how many of those times they:
a. Didn't make it in time to the toilet (rush)

b. Didn't feel the bowel movement until after it had happened (passive leakage)
The FI episodes will be calculated from those of 'Uncontrolled bowel Movements', whether this be 'rush' or 'passive leakage', by adding the two together.

DESCRIPTIVE ANALYSES

NB To help identify problems with missing data, outlying values, or other errors, full descriptive statistics MUST be produced for all variables in the database(s)

Participant flow

Participant throughput will be summarised in a CONSORT diagram.

Representativeness of sample
N/A

Baseline comparability of randomised groups

<u>Demographics</u>
Age and sex distributions will be described, by trial arm

<u>Prior and concurrent medications</u>
Proportion of patients taking loperamide and/or codeine

<u>Baseline and screening conditions</u>
Severity of FI symptoms according to bowel diary (i.e. mean number of episodes per week, recorded over a 14 day period)

Baseline medical history
Previous treatment for FI, and obstetric history. These are binary variables, except for number of previous vaginal deliveries (0, 1, 2, 3, 4, 5 or 6), and will be reported as number and percentage.

Baseline physical exam
N/A

Cluster characteristics if cluster randomised
N/A

Characteristics of care providers where applicable
N/A

Comparison of losses to follow-up

The proportion of patients withdrawing will be compared descriptively by arm

Comparison of compliance to treatment and protocol

The proportion of patients attending one or more treatment sessions will be compared descriptively.
The distribution of the proportion of treatment sessions attended per patient will be compared descriptively by arm. Participants who receive ≥ 10 treatments in 13 weeks will be considered to have received a full set of treatments for the per protocol analysis.

Emergency or accidental unblinding of randomised treatment

It was hard to envisage any necessity to break the randomisation code. We specified that should this be required, in the first instance the permission of the Local Principal Investigator should be sought. If they were unavailable, or this was not possible, the Academic Clinical Fellow, Emma Horrocks, of the Chief Investigator, Charles Knowles should be contacted.
Once permission had been sought, the local investigator could break the randomisation code by looking at CRF 4 for the appropriate participant. If this was not possible, because the information was unavailable out of hours, the lead centre should be contacted. In the first instance the Trial Manager could be contacted, who could break the randomisation code by the computer programme, and if she was unavailable the Daniel Simpkins at Nottingham Clinical Trials Unit should be contacted. Only the trial manager, independent statistician and Nottingham representative had access to the randomisation data within the database.

INTERIM ANALYSES AND SAFETY MONITORING ANALYSES

Purpose of interim analyses

None were planned

ANALYSIS OF PRIMARY OUTCOME

Definition of outcome measure

Responder vs. non-responder: Defined as a 50% or greater reduction in FI episodes per week, comparing end of intervention bowel diary with baseline). i.e.
If 100% × (#FI(baseline) - #FI(end of intervention))/#FI(baseline) ≥ 50%, class as responder
Where #FI stands for the average number of FI episodes per week. Where %change is negative, this represents an increase in FI episodes; where it is positive, this represents a decrease in FI episodes.

Descriptive statistics for outcome measure

Number and proportion of responders (as defined above) in each arm

Primary analysis

A logistic regression, adjusting for mean number of FI episodes at baseline, with a fixed effect for treatment arm, and sex, will be fitted using the Stata command *xtmelogit*, specifying study centre as a random effect

Assumption checks and actions to be taken if they do not hold
None

Other analysis supporting the primary (inc. sensitivity analyses)
N/A

ANALYSIS OF SECONDARY OUTCOMES

Definition of outcome measures
Percentage reduction in FI episodes per week for three additional cut offs: an improvement of ≥ 25% vs. less, an improvement of ≥ 75% vs. less and an improvement of 100% vs. less

Mean reduction in FI episodes per week (continuous); i.e. average number of FI episodes per week for the two weeks post-treatment, minus average number of FI episodes per week at baseline (from pre-treatment two-week bowel diary)

Mean reduction in (a) uncontrolled rush FI episodes per week and (b) uncontrolled passive leakage FI episodes i.e. average number of FI episodes per week for the two weeks post-treatment, minus average number of FI episodes per week at baseline (from pre-treatment two-week bowel diary)

Patient centred outcomes (continuous)

St Mark's score (continuous)

FI-QOL (continuous) [the four domains will be handled as outcomes in four separate models]

EQ-5D (continuous)

SF-36 (continuous) [the eight domains will be handled as outcomes in eight separate models]

Likert scale of patient's global impression of success (0-10). (continuous)

Short urinary symptom assessment (ordered categorical)

Descriptive statistics for outcome measure
Mean and SD for symmetric continuous variables
Median, 10^{th} & 90^{th} centiles for skewed continuous variables
Number and % for binary and other categorical variables

Secondary analysis
Continuous outcomes will be modelled using a mixed effects linear regression with the command *xtmixed*, adjusting for the baseline level of the outcome, sex and including a random effect for study centre. It should be noted that for outcome (2), analysis of change in FI episodes as the outcome is less efficient than ANCOVA i.e. end of intervention mean number of episodes as outcome, adjusted for baseline mean number of episodes. In this case the outcome variable will be mean number of FI episodes per week at end of treatment adjusting for the covariate, baseline measure of the outcome. Binary outcomes will be modelled using a mixed effects logistic regression *xtmelogit*, adjusting for the baseline measure of the outcome, sex and including a random effect for study centre. Urinary symptoms (outcome 9) will not be modelled.

Assumption checks and actions to be taken if assumptions do not hold.
Normality and homoscedasticity of residuals (*xtmixed* only), normality of random effects

Other analysis supporting the secondary (inc. sensitivity analyses)
N/A

SAFETY AND TOLERABILITY ANALYSES

Intervention exposure
N/A

All adverse events
The PTNS treatment and sham have no recognised significant adverse effects. However, at each weekly visit each patient will be asked if they have suffered any side effects or adverse effects of the treatment. These will be documented and in the study database and reported to the data monitoring committee prior to each meeting with them.

Adverse events leading to withdrawal

Serious adverse events

Clinical laboratory evaluations

SUBGROUP ANALYSES

Definition of outcome measure

Definition of subgroups

Primary outcome –
Sex (male vs. female)
FI severity (< or ≥ 7 episodes/wk)

Secondary outcomes –
age (<40 years, 40 to 60 years, 60+ years),

Sample size justification for the subgroup analysis
None

Descriptive analysis for subgroups

Method of analysis

An interaction term will be defined by multiplying the sub group dummy variables by the treatment assignment variable. For age, a global test of the two interaction terms will be performed using a likelihood ratio test.

AMENDMENTS TO VERSION 1.0

The pre-planned subgroup analysis 'urge versus passive FI episodes' had been inadvertently left out of version 1.0.

- Further, the sub-group analyses were to be conducted only on the primary outcome and not on the secondary outcomes as previously written.

- The Stata code listed in appendix 14(a) was refined during the analysis to correctly handle patients with fewer than 7 days of bowel diary data and the new code has replaced that previously in 14(a). The code to define stool consistency outcomes was not included in version 1.0 and was developed during the analysis.

- The listing of variables used in the multilevel multiple imputations was updated during the analysis to include more detail and also the imputation model for the four binary outcomes (primary (>= 50% reduction in FI) and three secondary (>=25%, >=75%, 100%)).

REFERENCES

Kahan BC & Morris TP. Analysis of multicentre trials with continuous outcomes: when and how should we account for centre effects? Statist. Med 2013, 32 1136–1149

Vaizey CJ, Carapeti E, Cahill JA, et al. Prospective comparison of faecal incontinence grading systems *Gut* 1999 44: 77-80 doi: 10.1136/gut.44.1.77

APPENDICES

Stata code for creating outcome variables

***** Bowel diary data**

```
forvalues i = 1(1)14 {
replace Q3_`i' = "" if Q3_`i' == "."
encode Q3_`i', generate(x)
drop Q3_`i'
rename x Q3_`i'
recode Q3_`i' (1=0)(2=1)
label variable Q3_`i' YN

replace Q4a_`i'="" if Q4a_`i' == "."
encode Q4a_`i', generate(x)
drop Q4a_`i'
rename x Q4a_`i'
recode Q4a_`i' (1=0)(2=1)
label variable Q4a_`i' YN

replace Q4b_`i'="" if Q4b_`i' == "."
encode Q4b_`i', generate(x)
drop Q4b_`i'
rencode x Q4b_`i'
recode Q4b_`i' (1=0)(2=1)
label variable Q4b_`i' YN

replace Q5_`i'="" if Q5_`i' == "."
encode Q5_`i', generate(x)
drop Q5_`i'
rencode x Q5_`i'
recode Q5_`i' (1=0)(2=1)
label variable Q5_`i' YN

generate x=real(Q1_`i')
drop Q1_`i'
rencode x Q1_`i'
generate x=real(Q2a_`i')
drop Q2a_`i'
rencode x Q2a_`i'
generate x=real(Q2b_`i')
drop Q2b_`i'
rencode x Q2b_`i'

encode Q6_`i', generate(x)
drop Q6_`i'
rencode x Q6_`i'
recode Q6_`i' (1=.)(6=0)(5=1)(4=2)
label variable Q6_`i' consist
```

```
}
sort PIN
compress

egen controlled = rowtotal(Q1_1 Q1_2 Q1_3 Q1_4 Q1_5 Q1_6 Q1_7 Q1_8
Q1_9 Q1_10 Q1_11 Q1_12 Q1_13 Q1_14), missing
egen uncontrolled_a = rowtotal(Q2a_1 Q2a_2 Q2a_3 Q2a_4 Q2a_5 Q2a_6
Q2a_7 Q2a_8 Q2a_9 Q2a_10 Q2a_11 Q2a_12 Q2a_13 Q2a_14), missing
egen uncontrolled_b = rowtotal(Q2b_1 Q2b_2 Q2b_3 Q2b_4 Q2b_5 Q2b_6
Q2b_7 Q2b_8 Q2b_9 Q2b_10 Q2b_11 Q2b_12 Q2b_13 Q2b_14), missing

egen staining = rowtotal(Q3_1 Q3_2 Q3_3 Q3_4 Q3_5 Q3_6 Q3_7 Q3_8 Q3_9
Q3_10 Q3_11 Q3_12 Q3_13 Q3_14), missing
egen pads = rowtotal(Q4a_1 Q4a_2 Q4a_3 Q4a_4 Q4a_5 Q4a_6 Q4a_7 Q4a_8
Q4a_9 Q4a_10 Q4a_11 Q4a_12 Q4a_13 Q4a_14), missing
egen enema = rowtotal(Q4b_1 Q4b_2 Q4b_3 Q4b_4 Q4b_5 Q4b_6 Q4b_7 Q4b_8
Q4b_9 Q4b_10 Q4b_11 Q4b_12 Q4b_13 Q4b_14), missing
egen social = rowtotal(Q5_1 Q5_2 Q5_3 Q5_4 Q5_5 Q5_6 Q5_7 Q5_8 Q5_9
Q5_10 Q5_11 Q5_12 Q5_13 Q5_14), missing
egen stool = rowtotal(Q6_1 Q6_2 Q6_3 Q6_4 Q6_5 Q6_6 Q6_7 Q6_8 Q6_9
Q6_10 Q6_11 Q6_12 Q6_13 Q6_14), missing

egen m_unc_a = rowmiss(Q2a_1 Q2a_2 Q2a_3 Q2a_4 Q2a_5 Q2a_6 Q2a_7)
egen m_unc_b = rowmiss(Q2b_1 Q2b_2 Q2b_3 Q2b_4 Q2b_5 Q2b_6 Q2b_7)

generate FI_episodes = uncontrolled_a + uncontrolled_b
generate FI_epi_pw = FI_episodes/2


*** primary outcome

generate Responder_50 = 1 if FI_epi_pw2/FI_epi_pw0<=.5
replace Responder_50 = 0 if FI_epi_pw2/FI_epi_pw0>.5 &
!missing(FI_epi_pw2) & !missing(FI_epi_pw0)

*** binary secondary outcomes: 25%, 75%, 100% improvement
generate Responder_25 = 1 if FI_epi_pw2/FI_epi_pw0<=.75
replace Responder_25 = 0 if FI_epi_pw2/FI_epi_pw0>.75 &
!missing(FI_epi_pw2) & !missing(FI_epi_pw0)

generate Responder_75 = 1 if FI_epi_pw2/FI_epi_pw0<=.25
replace Responder_25 = 0 if FI_epi_pw2/FI_epi_pw0>.25 &
!missing(FI_epi_pw2) & !missing(FI_epi_pw0)

generate Responder_100 = 1 if FI_epi_pw2==0
replace Responder_100 = 0 if FI_epi_pw2>0 & !missing(FI_epi_pw2)


*** Gastrointestinal quality of life

renpfix x_

foreach var of varlist GI2 GI4 GI6 GI8 GI9 GI11 GI12 GI13 GI14 GI16
GI18 GI20 GI22 GI24 GI26 GI28 GI30 GI32 GI34 GI36 {
*** reverse scoring
g x_`var'=1 if `var'==5
replace x_`var'=2 if `var'==4
replace x_`var'=3 if `var'==3
replace x_`var'=4 if `var'==2
replace x_`var'=5 if `var'==1
drop `var'
```

```
}

renpfix x_
order GI1 GI2 GI3 GI4 GI5 GI6 GI7 GI8 GI9 GI10 GI11 GI12 GI13 GI14
GI15 GI16 GI17 GI18 GI19 GI20 GI21 GI22 GI23 GI24 GI25 GI26 GI27 GI28
GI29 ///
  GI30 GI31 GI32 GI33 GI34 GI35 GI36
egen GIQoL_tot=rowtotal(GI1-GI36), missing
egen mGIQoL = rowmiss(GI1-GI36)
replace GIQoL_tot =. if mGIQoL ~= 0
```

***** EQ-5D**

```
rename EQ1 mob
rename EQ2 self
rename EQ3 usual
rename EQ4 pain
rename EQ5 mood
rename EQ6 VAS

gen EuroQol = 1
replace EuroQol = 1-.069 if mob == 2
replace EuroQol = 1-.314 if mob == 3
replace EuroQol = EuroQol-.104 if self == 2
replace EuroQol = EuroQol-.214 if self == 3
replace EuroQol = EuroQol-.036 if usual == 2
replace EuroQol = EuroQol-.094 if usual == 3
replace EuroQol = EuroQol-.123 if pain == 2
replace EuroQol = EuroQol-.386 if pain == 3
replace EuroQol = EuroQol-.071 if mood == 2
replace EuroQol = EuroQol-.236 if mood == 3
replace EuroQol = EuroQol-.081 if (mob ~= 1 |self ~= 1|usual ~=
1|pain ~= 1|mood ~= 1)&(mob ~= . & self ~= . & usual ~= . & pain ~= .
& mood ~= .)
replace EuroQol = EuroQol - .269 if mob == 3|self ==3 |usual ==3
|pain == 3|mood == 3
replace EuroQol=. if mob == .|self == .|usual == .|pain == .|mood ==
.
egen itemEuro = rownonmiss(mob self usual pain mood)
generate invalidEuro = 1 if itemEuro > 0 & itemEuro < 5
replace invalidEuro = 0 if itemEuro == 5
egen mEuroQol = rowmiss(mob self usual pain mood)
```

***** Patient Centred Outcomes**

```
egen PC_tot=rowtotal(PC1 PC2 PC3 PC4 PC5 PC6 PC7 PC8), missing
egen PC_mean=rowmean(PC1 PC2 PC3 PC4 PC5 PC6 PC7 PC8)
egen mPC = rowmiss(PC1 PC2 PC3 PC4 PC5 PC6 PC7 PC8)
replace PC_mean=. if mPC ~= 0
```

***** St. Mark's FI scale**

```
recode CC1-CC4 (1=0)(2=1)(3=2)(4=3)(5=4)
recode CC5 CC6 (1=0)
recode CC7 (1=0)(2=4)
egen CC_tot=rowtotal(CC1 CC2 CC3 CC4 CC5 CC6 CC7), missing
egen mCC = rowmiss(CC1 CC2 CC3 CC4 CC5 CC6 CC7)
replace CC_tot = . if mCC ~= 0
```

```
*** SF-36
rename SF1 q1
rename SF2 q2
rename SF3a q3
rename SF3b q4
rename SF3c q5
rename SF3d q6
rename SF3e q7
rename SF3f q8
rename SF3g q9
rename SF3h q10
rename SF3i q11
rename SF3j q12
rename SF4a q13
rename SF4b q14
rename SF4c q15
rename SF4d q16
rename SF5a q17
rename SF5b q18
rename SF5c q19
rename SF6 q20
rename SF7 q21
rename SF8 q22
rename SF9a q23
rename SF9b q24
rename SF9c q25
rename SF9d q26
rename SF9e q27
rename SF9f q28
rename SF9g q29
rename SF9h q30
rename SF9i q31
rename SF10 q32
rename SF11a q33
rename SF11b q34
rename SF11c q35
rename SF11d q36

foreach var of varlist q1 q2 q20 q22 q34 q36 {
generate `var'_value=100 if `var'==1
replace `var'_value=75 if `var'==2
replace `var'_value=50 if `var'==3
replace `var'_value=25 if `var'==4
replace `var'_value=0 if `var'==5
}

foreach var of varlist q3-q12 {
generate `var'_value=0 if `var'==1
replace `var'_value=50 if `var'==2
replace `var'_value=100 if `var'==3
}

foreach var of varlist q13-q19 {
generate `var'_value=0 if `var'==1
replace `var'_value=100 if `var'==2
}

foreach var of varlist q21 q23 q26 q27 q30 {
generate `var'_value=100 if `var'==1
replace `var'_value=80 if `var'==2
replace `var'_value=60 if `var'==3
```

```
replace `var'_value=40 if `var'==4
replace `var'_value=20 if `var'==5
replace `var'_value=0 if `var'==6
}

foreach var of varlist q24 q25 q28 q29 q31 {
generate `var'_value=0 if `var'==1
replace `var'_value=20 if `var'==2
replace `var'_value=40 if `var'==3
replace `var'_value=60 if `var'==4
replace `var'_value=80 if `var'==5
replace `var'_value=100 if `var'==6
}

foreach var of varlist q32 q33 q35 {
generate `var'_value=0 if `var'==1
replace `var'_value=25 if `var'==2
replace `var'_value=50 if `var'==3
replace `var'_value=75 if `var'==4
replace `var'_value=100 if `var'==5
}

drop q1-q36

egen SF36_PF = rowmean(q3_value q4_value q5_value q6_value q7_value
q8_value q9_value q10_value q11_value q12_value)
egen SF36_RLPH = rowmean(q13_value q14_value q15_value q16_value)
egen SF36_RLEM = rowmean(q17_value q18_value q19_value)
egen SF36_EF = rowmean(q23_value q27_value q29_value q31_value)
egen SF36_EM = rowmean(q24_value q25_value q26_value q28_value
q30_value)
egen SF36_SF = rowmean(q20_value q32_value)
egen SF36_P = rowmean(q21_value q22_value)
egen SF36_GH = rowmean(q1_value q33_value q34_value q35_value
q36_value)
```

***** QoL FI scale**

```
generate FI1_rev=1 if FI1==5
replace FI1_rev=2 if FI1==4
replace FI1_rev=3 if FI1==3
replace FI1_rev=4 if FI1==2
replace FI1_rev=5 if FI1==1

egen FIQoL_lif = rowmean(FI2a FI2b FI2c FI2d FI2e FI2g FI2h FI3b FI3l
FI3m)
egen FIQoL_cop = rowmean(FI2f FI2i FI2j FI2k FI2m FI3c FI3h FI3j
FI3n)
egen FIQoL_dep = rowmean(FI1_rev FI3d FI3f FI3g FI3i FI3k FI4)
egen FIQoL_emb = rowmean(FI2l FI3a FI3e)

egen mFIQoL_lif = rowmiss(FI2a FI2b FI2c FI2d FI2e FI2g FI2h FI3b
FI3l FI3m)
egen mFIQoL_cop = rowmiss(FI2f FI2i FI2j FI2k FI2m FI3c FI3h FI3j
FI3n)
egen mFIQoL_dep = rowmiss(FI1_rev FI3d FI3f FI3g FI3i FI3k FI4)
egen mFIQoL_emb = rowmiss(FI2l FI3a FI3e)

replace FIQoL_lif =. if mFIQoL_lif ~= 0
replace FIQoL_cop =. if mFIQoL_cop ~= 0
```

```
replace FIQoL_dep =. if mFIQoL_dep ~= 0
replace FIQoL_emb =. if mFIQoL_emb ~= 0
```

Variables to be used in the multilevel multiple imputation

Outcomes that are reported as a group of domains will be imputed together where possible. Baseline measures, mid-treatment (bowel diary data only) and end of treatment outcomes will be included together in the multivariate response. Centre is the only level 2 variable (random intercepts)

	Imputation Variables	
Outcomes	To include in multivariate response	Auxillary variables
Mean number of FI episodes per week	Mean number of FI episodes per week, St Mark's Continence Score, Likert	Age, Sex and random allocation
GI QOL Index (reported as total score	Mean number of FI episodes per week, St Mark's Continence Score, GI QOL	Age, Sex and random allocation
Patient centered outcomes form	Patient centered outcomes form	Age, Sex and random allocation
SF-36 (eight domains)	SF-36 (4 domains at a time)	Age, Sex and random allocation
Quality of life scale for FI (four domains)	QOL for FI	Age, Sex and random allocation
St Marks Continence Score	Mean number of FI episodes per week, St Mark's Continence Score	Age, Sex and random allocation
EQ-5D	EQ-5D	Age, Sex and random allocation

List of case report forms (see statistics master file for detail)

☐ CRF 1 – Eligibility assessment

☐ CRF 2 – Initial assessment

☐ CRF 3 – PRE-TREATMENT questionnaires

☐ CRF 3 – POST-TREATMENT questionnaires

☐ CRF 4 – Randomisation

☐ CRF 5 – Record stimulation parameters adverse events and medication / pad usage

☐ CRF 6 – Post-treatment information

☐ CRF 7 – Final study visit information

☐ BOWEL DIARY – PRE-TREATMENT (14 days)

☐ BOWEL DIARY – MID-TREATMENT (7 days)

☐ BOWEL DIARY – POST-TREATMENT (14 days)