ELECTRIC

electric tibial nerve stimulation to reduce incontinence in care homes

Statistical Analysis Plan (version 1.0)

A UK Collaborative Study funded by the
NIHR Evaluation, Trials and Studies Coordinating Centre (NETSCC),
Health Technology Assessment (HTA) Programme

FUNDED BY
NIHR | National Institute for Health Research

GCU Glasgow Caledonian University
CHaRT Centre for Healthcare Randomised Trials

1 Administrative information

This SAP is based, as far as is appropriate, on guidelines given in JAMA. 2017;318(23):2337-2343. doi:10.1001/jama.2017.18556

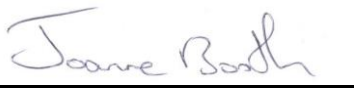
TRIAL FULL TITLE	Electric Tibial Nerve Stimulation to Reduce Incontinence in Care Homes
EUDRACT NUMBER	n/a
SAP VERSION	Draft version 1 (based on Protocol ELECTRIC (V3.0 xx/xx/xx))
ISRCTN NUMBER	98415244 HTA Project: 15/130/73
CLINICAL TRIALS.GOV NUMBER	NCT03248362
SAP VERSION DATE	Sept 2018
TRIAL STATISTICIAN	Dr Lorna Aucott
TRIAL CHIEF INVESTIGATOR	Professor Joanne Booth
SAP AUTHOR	Dr Lorna Aucott (Senior Statistician CHaRT)

1.1 SAP Signatures

I give my approval for the attached SAP entitled ELECTRIC, dated 25th October 2019

Chief Investigator

Name:

Signature: 

Date: 5th November 2019

Statistician

Name:

Signature: 

Date: 5th November 2019

1.2 Table of Contents

1.3 List of Tables

1.4 List of Figures

1.5 Appendices

1.6 Abbreviations and Definitions

AE	Adverse Event
CH	Care Home
CHaRT	Centre for Healthcare Randomised Trials
CI	Chief Investigator
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
DMEC	Data Monitoring and Ethics Committee
FI	Faecal Incontinence
GP	General Practitioner
HSRU	Health Services Research Unit
HTA	Health Technology Assessment
ISRCTN	International Standard Randomised Controlled Trial Number
MMSE	Mini Mental State Examination
MTSQ	Minnesota Toileting Skills Questionnaire
NHS	National Health Service
NIHR	National Institute Health Research
NRES	National Research Ethics Service
PI	Principal Investigator
PMG	Project Management Group
PTNS	Percutaneous Posterior Tibial Nerve Stimulation
PVRU	Post Void Residual Urine volume
PWT	Pad Weight Test
QoL	Quality of Life
RCT	Randomised Controlled Trial
RUQ	Resource Use Questionnaire
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SC	Senior Carer
SD	Standard Deviation
TPTNS	Transcutaneous Posterior Tibial Nerve Stimulation
TSC	Trial Steering Committee
UK	United Kingdom
UKCRC	United Kingdom Clinical Research Collaboration
UI	Urinary Incontinence
UoA	University of Aberdeen

2 Introduction

2.1 Background

The highest prevalence of urinary incontinence (UI), defined by the International Continence Society as 'any urinary leakage', is found in residential or nursing care homes (CH). UI is distressing for older adults and profoundly impacts on dignity and quality of life. It is associated with impaired physical functioning cognitive impairment sleep disturbance, falls, fractures, hygiene and tissue viability problems. UI affects participation by older adults and is a major cause of clinical depression and social isolation. Incontinence is costly, to CH providers, the NHS and the individual older adult. Direct personal and treatment costs are high. Intangible costs associated with social isolation and withdrawal from participatory groups also occur but have not been quantified.

2.2 Rationale of the analyses

To assess if a programme of transcutaneous posterior tibial nerve stimulation (TPTNS) is a clinically effective treatment for urinary incontinence (UI) in care home residents and the associated costs and consequences. The hypothesis being tested is that residents receiving TPTNS find it beneficial to the volume of urinary incontinence (UI) compared to the residents receiving direct Sham stimulation.

2.3 Study Objectives

A. To establish whether TPTNS is more effective than sham stimulation for reducing the volume of urinary incontinence at 6, 12 and 18 weeks, in care home residents.

B. To investigate mediating factors that impact on the effectiveness of TPTNS in a mixed method, process evaluation involving fidelity, implementation support and qualitative components.

C. To undertake economic evaluation of TPTNS in care homes assessing the costs of providing the programme and presenting them alongside the key primary and secondary outcomes in a cost-consequence analysis.

D. To explore in an interview study the experiences of TPTNS from the perspectives of:

    a. Care home residents

    b. Family carers

    c. Care home nurses and senior carers

    d. Care home managers

This SAP covers objectives A and B.

3 Study Methods

3.1 Trial design

ELECTRIC is a pragmatic multicentre 2-arm randomised controlled superiority trial, to compare effectiveness of TPTNS (n=250) with sham stimulation (n=250) as a control, to reduce UI in CH residents. Participants and outcomes assessors will be blinded, but none of the healthcare providers who administer the treatment or sham). Outcomes are assessed at 6, 12 and 18 weeks post randomisation – detailed later for each measure in section 3.6.3.

3.2 Randomisation and Blinding

Eligible and consenting participants will be randomised to one of the two groups. This will be by a web-based application hosted by the fully registered with the UK Clinical Research Collaboration (UKCRC), Clinical Trials Unit (CTU) at the Centre for Healthcare Randomised Trials (CHaRT), Health Services Research Unit (HSRU) in Aberdeen.

Randomisation will be computer allocated based on random permuted blocks of size two, four or six, stratified by:

- Sex – male/female
- UI severity - mild (0-200ml/24 hrs); moderate (200-400ml /24 hrs); severe (400+ ml/24 hrs)
- Centre.

3.3 Sample Size and Power Calculation

An amendment to the final sample size requirement has been required. Recruitment has not followed expected projections. The original recruitment target (see ELECTRIC protocol v1.0) was calculated to be 500 care home residents needed to detect differences of 200ml/24hrs with 90% power at the two-sided 5% alpha level. (This is 344 but inflated by an attrition estimate of 30% to account for loss due to death and other types of loss to follow up). The standard deviation (SD) used in the original calculation came from a single-centre trial, with a selected population where the reported SD was 450ml. A 95% CI was put around the SD estimate and used for the sample size calculation to account for recruiting to a pragmatic multi-centre trial.

Kieser and Friede [13] recommend re-estimating the sample variance from observed data using the whole trial cohort and calculating the one-sample variance, and also an adjusted estimate to account for potential bias in the one-sample variance under the alternative hypothesis. Following one full year of recruitment to the ELECTRIC trial, a data cut was performed and the sample size reviewed by an independent statistician – see Appendix B. This indicated attrition to be closer to 15% than 30% and that the SD for the primary outcome was estimated to be less than 450ml. This report concluded that a sample size of 278 would satisfy power and difference requirements. However, based on the original power calculation (n=344) and allowing a reduced attrition of 15% suggests instead that an overall sample size 405 would be sufficient. The PMG discussed the findings and concluded that recruitment should continue for the length of the planned recruitment period (18 months) with the aim of exceeding the minimum requirement of 278 randomised participants. Three blinded, independent statisticians, the DMEC, TSC, PMG and funders all agreed with this sample-size revision. By that stage the internal team had set a target of 400 by the end of month 25 (July 2019)

3.4 Interim Analyses and Data Monitoring

There are no planned interim analyses for the ELECTRIC Trial. Trial oversight committees (DMC and TSC) will however receive detailed reports on trial progress and safety data at least annually.

3.4.1 Documentation of Report Summaries

The data available for each report will be preserved, along with all documentation of analysis plans, programming code and reporting provided.

3.5 Timing of final Analyses

The final analyses will be performed when sufficient numbers have completed their 6-week assessment (time of the primary outcome) or have dropped out prior to the 6 week assessment. This will be 344 (the minimum number required by the power calculation +15% to allow for drop out for the follow-up period).

3.6 Outcome Measurements

The outcome measurements will be taken within a one-week period at defined times (6, 12 and 18 weeks) post randomisation.

3.6.1 Primary Outcome

- Volume of UI over 24 hour period at 6 weeks measured by 24 hour PWT

3.6.2 Secondary Outcomes (at 6, 12 and 18weeks unless otherwise stated)

- 24 hour PWT at 12 and 18 weeks to assess sustainability
- Number of pads used in 24 hours
- Post-void residual urine volume (PVRU)
- Resident, family carers, staff perception of bladder condition (P-PBC, FC-PBC, S-PBC)
- Resident Toileting Skills (Minnesota Toileting Skills Questionnaire- MTSQ), assessed by Residents and Staff
- Quality of life - DEMQOL (self-reported by the patient) and DEMQOL-Proxy (proxy reported by a carer)

3.6.3 Timing of Outcome Measures

Table 3-1: Outcome timings

	Baseline	6-week	12-weeks	18weeks
24 hour PWT	●	●	●	●
Number of pads used	●	●	●	●
24 hour bladder diary	●	●	●	●
PVRU *	●	●	●	●
P-PBC *	●	●	●	●
FC-PBC *	●	●	●	●
S-PBC *	●	●	●	●
MTSQ-R *	●	●	●	●
MTSQ-S *	●	●	●	●
DEMQOL *	●	●		●
DEMQOL-proxy *	●	●		●

* Scales and Sub-scores derived from participant questionnaires

In addition to these outcomes measures, adherence and fidelity of the treatment administration was monitored over the intervention period. For adherence the target for all

residents was to receive at least 8 of the 12 sessions by 6 weeks post randomisation the time required for the full stimulation program. With respect to fidelity the target was for all participating residents to receive the intervention duration and intensity of stimulation and correct ankle position as per the protocol associated with the group to which they were allocated. See section 6.1 for how this will be incorporated into the analyses.

4 Statistical principles

4.1 Summaries, Estimates, Confidence intervals and p-values

Statistical analysis will be tested at the 2-sided 5% significance level with any estimates displayed with 95% confidence intervals (CIs) and p-values.
Any p-values ≥0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001". The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

4.2 Adherence and Protocol deviations

Originally this trial had a Stop/Go criteria based on results from an internal pilot with 100-140 residents to determine progression to full trial. The criteria are based on recruitment, adherence to stimulation, completeness of the primary outcome measure and fidelity of the stimulations - see Protocol V3 for Stop/go full details. These targets were met and no protocol deviations were required.

4.3 Analysis populations

Statistical analysis will be intention-to-treat based on all subjects who were randomised. It is not expected that any participant will switch treatments.
The Safety population will consist of participants who receive any trial intervention. Participants not receiving an intervention will be excluded from safety analysis.

5 Trial Population

The CONSORT flow diagram (fig 1) describes the participant flow through the trial. The level of withdrawals either from the interventions and/or from follow up will be monitored along with any details and/or reasons given. The timing of any withdrawals will be summarised.

6 Statistical Analysis

For each treatment group the general baseline characteristics will be descriptively summarised (See Table 7-2). The other measures will also be summarised by treatment groups at baseline (Tables 7-3 to 7-5) but additionally at each follow-up timepoint within the analyses tables (Tables 7-7 to 7-10). These will be as sample size, means (with standard deviations), medians (with inter-quartile ranges) and minimum and maximum or counts (with percentages) where appropriate.

Further analyses will be conducted on all the primary and secondary outcomes to assess treatment effects. These will be performed on the basis of the intention-to-treat principle.

6.1 Primary Outcome

The primary outcome, measured at 6 weeks post randomisation, total volume of urine leaked in 24 hours, will be analysed using linear regression correcting for baseline 24-hours PWT

and the design covariates (severity (mild/moderate/severe) and gender (m/f). Any possible care home clustering will be accounted for using a random effects robust variance.

In addition, using definitions given at the end of section 3.6.3, we will explore effects of adherence to treatment and fidelity, using randomisation as the instrumental variable in a complier average causal effect (CACE) model,[12] (using 2-stage least squares).

6.2 Secondary Outcomes

Secondary outcomes will be analysed using a similar strategy but with models suitable for each outcome. In addition these will utilise all available follow-up data from all randomised participants using a standard time interaction model used to incorporate the repeated measures. These will be estimated using GLM linear regression for continuous data, since although outcomes may be skewed the use of baseline information as a covariate will satisfy the normal assumptions. For binary outcomes we will use Poisson regression models with a log link function summarising the treatment effects as adjusted risk differences (ADR) and as adjusted relative risk (ARR) ratios. All models will be adjusted as described above.

All model assumptions will be assessed by means of the summary statistics and/or graphical plots.

6.3 Derived Variables - Patient Reported Outcome Measures (PROMs):

There are a number of PRO trial data collected using validated questionnaires, which then are combined into an overall score. These are indicated by * in Table 3-1. Codes for these are developed in-house, checked and then validated with dummy data by an independent statistician. For any amalgamated scores missing items will be imputed using the strategies as set out and agreed by the project team in Appendix B, taking into account the level of missingness overall and within a person.

6.4 Subgroup Analyses

Subgroup will be tested using interactions. These will be for:

- Gender (Male/ Female)
- UI severity (mild (0-200ml/24 hrs); moderate (>200-400ml /24 hrs); severe (>400ml/24 hrs))
- Functional Dependency
    - Total Barthel
    - Barthel Mobility 2.4 (4 groups)
    - Barthel Toilet 2.7 (3 groups)
- Clinical frailty scale – (<=5, 6, 7 or more)
- On anti-cholinergics for incontinence (or not)
- Falls status in last 6 months
    - Number residents who at baseline have fallen in the last 6 months [n (%)]
      No falls: <=6 falls : >6

The pre-defined subgroups will be reported as the magnitude of the subgroup effect estimates along with their 95%CI's. These will be interpreted in an exploratory manner and interpreted broadly and thus provide recommendations for further investigations.
Post- hoc: Forest plots will be used to illustrate possible effect of:

- the size or location (urban or rural) of the centre may be useful
- Plus Split into full 12 session vs rest

6.5 Missing Data

The extent of missing data will be reported explicitly for the key primary and secondary endpoints. It is not expected for there to be extensive missingness in ELECTRIC. Only if this is > 10% and/or known to be not completely at random will this be considered. Should this be the case then, missing data for the primary outcome only will be handled using appropriate methods[6], probably using multiple imputation methods, but will depend on the amount and pattern of missingness. We will also conduct a sensitivity analysis to test assumptions[7].

6.6 Technical Details

The current Protocol (version 3 at time of writing) will be consulted for this SAP. All statistical analyses will be conducted using Stata (version 15 at time of writing). All results will be processed into a PDF/Word directly from Stata via LaTEX (MiKTeX 2.9 at time of writing)

A second review statistician will independently reproduce the primary analyses and some random selected summary statistics tables. The reviewing statistician will have an overview of the entire analyses and will explicitly check the code producing tables (selected at random) as well as any other pieces of code as desired.

6.7 Prior and Concurrent Medications and Medical Conditions

Other than the anti-cholinergics for incontinence which will be examined as the sub-group analyses any other previous/concurrent medication will be varied and should not act as a moderator to the intervention and hence will be treated pragmatically as a random variable across the arms of the trial.

6.8 Adverse Events Analyses

An adverse event (AE) is defined as any untoward medical occurrence in a participant, not necessarily having a causal relationship. Non-serious adverse events will not be collected or reported. An adverse event is defined as "serious" (SAE) if it

- Results in death
- Is life threatening
- Requires or prolongs inpatient hospitalisation
- Results in persistent/significant disability/incapacity
- Is otherwise considered medically significant by the investigator.

In the ELECTRIC trial the following related minor AEs are potentially expected:

- Transient skin redness at electrodes sites
- Minor itch at electrode sites

There are no related serious AEs expected in this trial given the previous established safety profile of the TPTNS, However, any serious related AEs that do occur will be recorded as such.

Hospitalisations for elective treatment of a pre-existing condition are not considered as an AE or SAE. Complications occurring during such hospitalisation are also not AEs or SAEs.

Falls, fractures, UTIs, emergency admissions to hospital in 6 month period prior to the resident's participation in the ELECTRIC trial are recorded at study baseline. They are not considered as related AEs/SAEs during the trial period unless they occur during a stimulation session or during the measurement of study outcomes.

7 Tables

7.1 Baseline

Table 7-1 Description of centres

Centre Name	Sham n=	TPTNS n=	Overall	Type of care	No Beds	Country
Total						
	Average cluster size Mean(sd)			% of Residential Nursing Other	Average number of beds Mean(sd)	% England Scotland

Table 7-2: General baseline characteristics

Characteristics	Scotland TPTNS N = xxx	Scotland Sham N =xxx	England TPTNS N =xxx	England Sham N =xxx
Age – n, mean (sd)*				
Female – n (%)				
Missing				
Mini Mental State Exam (Total Score) *				
Length of Stay (weeks): *				
DEMQoL –(Total Score) *				
DEMQoL – proxy (Total Score) *				
Barthel Score (0-20) *				
Clinical Frailty – n(%) very fit well managing Well vulnerable mildly frail moderately frail severely frail very severely frail terminally ill				

* n; mean(sd) median(IQR) and min/max

Table 7-3: Baseline Severity of UI

	TPTNS n=xxx	Sham n=xxx
Duration of UI (years/months) n, mean(sd)*		
Base UI severity – n(%)		
mild (0-200ml/24 hrs)		
moderate (>200-400ml /24 hrs)		
severe (>400ml/24 hrs)		
Total volume urine leaked in 24 hrs (ml/24 hrs) – n, mean(sd)*		
Number of pads used in 24 hrs – n, mean(sd)*		
ICIQ-SF		
How often does urine leak Never About once a week or less often Two or three times a week About once a day Several times a day All the time		
How much leaks None Small amount Moderate amount Large amount		
How much does leaking urine interfere (0-10) median(IQR)		
OVERALL ICIQ-SF Severity Score (0-21) n, mean(sd)* **Severity categories** None (0) Slight (1 – 5) Moderate (6 – 12) Severe (13 – 18) Very Severe (19 – 21)		
Type of UI - n(%) Urgency UI Stress UI Mixed UI Functional UI Obstructive UI Not recorded		
Is the resident recorded as BPH Y n (% men only)		
Wears Pads continuously (Y) – n(%)		
PVRU (ml) – n, mean(sd)*		
Recorded UTIs treated with antibiotics n, mean(sd)*		

* n, mean(sd); median(IQR); and min/max

Table 7-4: Baseline Function and Management of incontinence with falls, fractures & ulcer risk

Incontinence	TPTNS n=xxx	Sham n=xxx
UI Mobility limitation - n(%) multiple allowed		
Confined to bed		
Able to sit out but unable to stand unassisted		
Able to stand unassisted		
Mobile with assistance of ONE persons		
Mobile with assistance of TWO people		
Mobile with assistance of equipment eg walking frame, wheelchair		
Independently mobile around home		
Independently mobile out of the home		
Toilet access restrictions - n(%) multiple allowed		
Mobility		
Problems communicating their need to use the toilet		
Problems finding/locating toilet		
Does not try to get to the toilet		
Other		
Medication for Incontinence (Y) – n(%)		
MTSQ-R Resident (total score, 0-20) – n, mean(sd) *		
MTSQ-S – Staff (total score, 0-20) – n, mean(sd)*		
Falls (In the Previous 6 months):		
i. What was the:		
• Number residents who have fallen [n (%)]		
• Number of falls per resident [mean (sd)] *		
• Annual home falls rate [&]		
ii. Injurious falls		
iii. Recorded fracture(s)		
iv. Site(s) of recorded fracture(s) eg neck of femur, humerus		
v. Admissions or unplanned visits to A&E/hospital		
vi. Urology /urogyneacology/continence service appointments (may be in care home)		
vii. Pressure ulcers recorded in past month		
viii. Site(s) of recorded pressure ulcer(s)		
ix. Non-pressure wounds recorded in the past month		
Pressure ulcer risk score in past month (Y) – n(%)		

* n, mean(sd) median(IQR) and min/max
[$] All falls in home *2 /(number of residents)

Table 7-5: **Baseline for Bladder Condition**

Bladder Condition- n(%)	TPTNS n=xxx	Sham n=xxx
Patient Perception (P-PBC)		
no problems at all.		
some very minor problems.		
minor problems.		
(some) moderate problems.		
severe problems.		
many severe problems		
missing		
Carer Perception (FC-PBC)		
No problems at all.		
some very minor problems.		
minor problems.		
(some) moderate problems.		
severe problems.		
many severe problems		
missing		
Staff Perception (S-PBC)		
No problems at all.		
some very minor problems.		
minor problems.		
(some) moderate problems.		
severe problems.		
many severe problems		
missing		

7.2 Serious adverse Events

Table 7-6: Serious adverse Events

Adverse Events n(%)	**TPTNS**		**Sham**		**Total**	
People						
Male						
Female						
AEs						
Type of Adverse Event	**SAE**	**Non SAE**	**SAE**	**Non SAE**	**SAE**	**Non SAE**
Expected						
Related to Study Procedure						
Death	**Related**	**Not**	**Related**	**Not**	**Related**	**Not**

7.3 Outcome summaries and model Estimates

Table 7-7: Primary and Secondary outcome for Total volume urine leaked in 24 hrs outcome: Summaries*# and Model results

Urine leaked	TPTNS	Sham	Effect size	95% CI	p-value
Total volume urine leaked in 24 hrs (ml/24 hrs)			[a]		
• Baseline * • 6 weeks *(primary) • 12 weeks *(sustainability) • 18 weeks *(sustainability)					
Compliance #			[b]		

*Continuous data: n; mean (sd), median (IQR) and (min, max)
Binary x/n (%)
[a] Mean difference between TPTNS and SHAM
[b] CACE of TPTNS relative to SHAM

Table 7-8: Secondary outcomes: Number of pads used and PVRU: Summaries*# and Model results

	TPTNS	Sham	Effect size [a]	95% CI	p-value
Number of pads used in 24 hrs (count) # • Baseline • 6 weeks • 12 weeks • 18 weeks					
PVRU (ml) * • Baseline • 6 weeks • 12 weeks • 18 weeks					

Treated as pseudo continuous data: n, mean (sd), median(IQR) and (min, max)
*Continuous data: n, mean (sd) and median(IQR) (min, max)
[a] Mean difference between TPTNS and SHAM

Table 7-9: Secondary outcomes: Bladder Condition: Summaries* and Model results

Bladder Condition	TPTNS	Sham	Effect size [a]	95% CI	p-value
Patient Perception (P-PBC) * • Baseline • 6 weeks • 12 weeks • 18 weeks					
Carer Perception (FC-PBC) * • Baseline • 6 weeks • 12 weeks • 18 weeks					
Staff Perception (S-PBC) * • Baseline • 6 weeks • 12 weeks • 18 weeks					

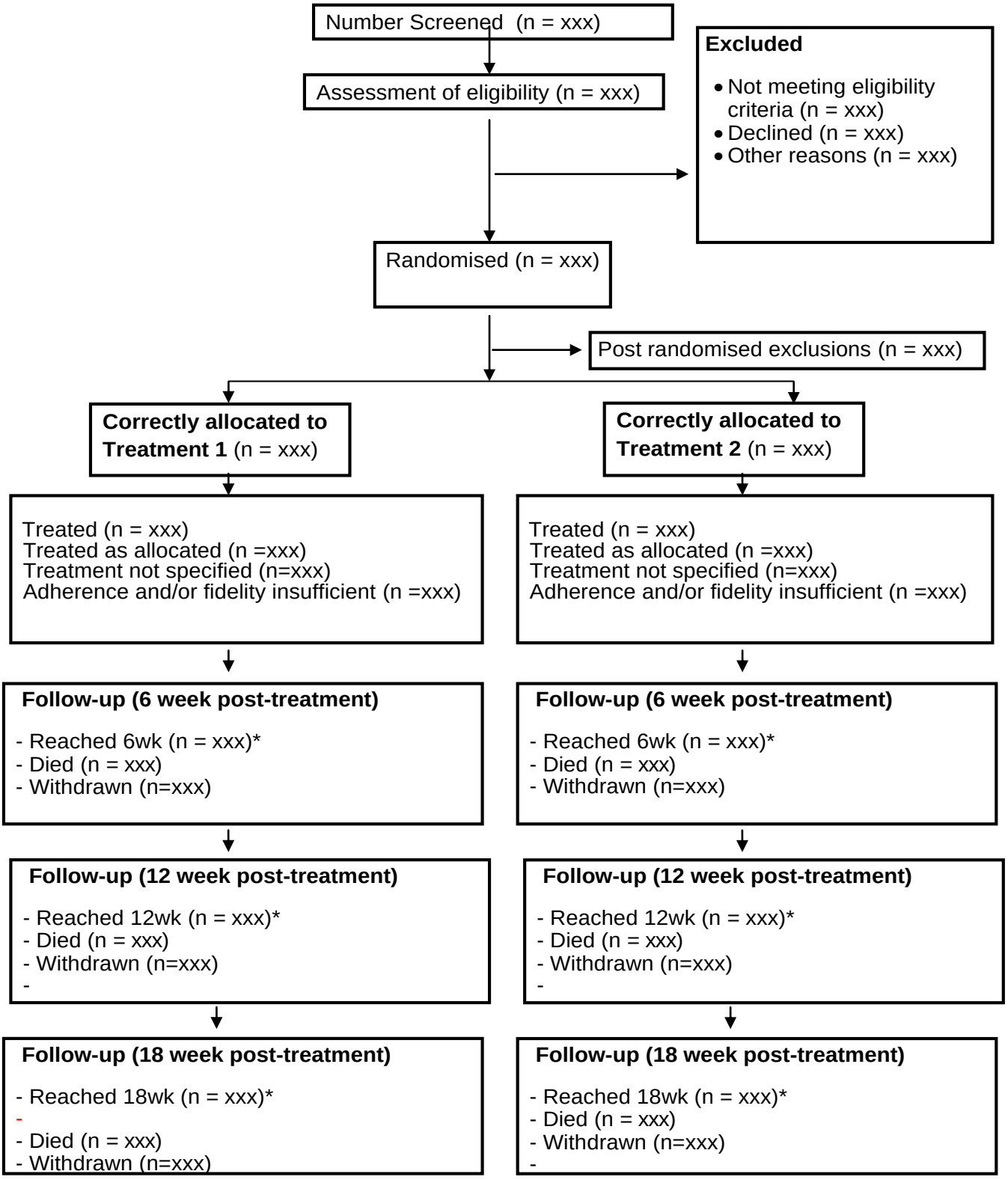
* Ordinal scale: but for modelling purposes treated as interval: n, mean (sd), median (IQR) and (min, max)
[a] Mean difference between TPTNS and SHAM

Table 7-10: Secondary Outcomes: Toilet Skills and QoL: Summaries and Model results

Toilet Skills and QoL	TPTNS	Sham	Effect size [a, b]	95% CI	p-value
MTSQ-R Resident (total score) (continuous) * • Baseline • 6 weeks • 12 weeks • 18 weeks					
MTSQ-R – Staff (total score) (continuous) * • Baseline • 6 weeks • 12 weeks • 18 weeks					
DEMQoL – (total Score) (continuous) * • Baseline • 6 weeks • 12 weeks • 18 weeks					
DEMQoL - proxy – (total Score) (continuous) * • Baseline • 6 weeks • 12 weeks • 18 weeks					

* Continuous scale: n, mean(sd), median(IQR) and (min, max)
[a] Adjusted for outcome at baseline, severity and gender with a random effect for care homes.
[b] Mean difference between TPTNS and SHAM

Figure 1 CONSORT Trial Flow Diagram



* included for monitoring purposes only during the course of the trial

8 References

1. Mahoney FI, Barthel D. Functional evaluation: the Barthel Index *Maryland State Medical Journal* 1965:14:56-61.
2. Smith SC, Lamping DL, Banerjee S. Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. *Health Technol Assess.* 2005 Mar;9(10):1-93.
3. Talley KM et al. Reliability and validity of two measures of toileting skills in frail older women without dementia. *J Gerontol Nurs* 2016:3:1–5. (- Minnesota Toilet Skills Questionnaire (MTSQ_R &S))
4. Coyne KS et al. The validation of the patient perception of bladder condition (PPBC): a single-item global measure for patients with overactive bladder. *Eur Urol.* 2006 Jun;49(6):1079-86.
5. Vinsnes A et al. Effect of physical training on urinary incontinence: a randomized parallel group trial in nursing homes. *Clin. Interv Ageing* 2012; 7: 45–50
6. Carpenter J, Kenward M. Missing data in clinical trials — a practical guide. In: Research NIfH, editor. Birmingham, 2008.
7. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011: 342:d40.
8. Greenland S. An introduction to instrumental variables for epidemiologists*. Int J Epidemiol.* 2000 Aug;29(4):722-9.
9. Nico Nagelkerke, Vaclav Fidler, Roos Bernsen and Martien Borgdor. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. *Statistics in Medicine* 2000; 19:1849-1864
10. Folstein, M., Folstein, S.E., McHugh, P.R. “Mini-Mental State” a Practical Method for Grading the Cognitive State of Patients for the Clinician. *Journal of Psychiatric Research*, 1975; 12(3): 189-198.
11. White IR and Thompson SG Adjusting for partially missing baseline measurements in randomised trials. *Statistics in Medicine*, 2005;24: 993-1007
12. G Dunn, R. A. Emsley, H Liu, S Landau, J Green, I White, and A Pickles. Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health. *Health Technology Assessment*, 2015;19(93)
13. Kieser M1, Friede T. Simple procedures for blinded sample size adjustment that do not affect the type I error rate. *Stat Med.* 2003 Dec 15;22(23):3571-81.

Appendix A: Supplemental Baseline Severity of UI

Severity of UI	**TPTNS =**	**Sham =**
Aware of need to open bowels (Y) - n(%)		
Aware of need to pass Urine (Y) - n(%)		
UI assessment recorded (Y) - n(%)		
FI assessment recorded (Y) - n(%)		
Assessment of both recorded (Y) - n(%)		
Duration of UI (years/months) mean(sd)*		
Duration FI (years/months) mean(sd)*		
Duration Of both FI and UI (years/months) mean(sd)*		
When does urine leak? - n(%) multiple allowed		
Never – urine does not leak		
Leaks before they can get to the toilet		
Leaks when they cough or sneeze		
Leaks when they are asleep		
Leaks when they are physically active/exercising		
Leaks when they have finished urinating and are dressed		
Leaks for no obvious reason		
Leaks all the time		
Bladder Symptoms - n(%)		
Urgency		
Frequency		
Nocturia		
Hesitancy		
Intermittency		
Post-micturition dribble		
Incomplete emptying		
Bladder pain		
Symptoms not recorded		

Appendix B: ELECTRIC missing value criteria for PROMS

Tool	References	Scoring	Indication	Missing data
DEMQOL 'DEMQOL and DEMQOL-Proxy are intended for use in evaluating HRQL in group comparisons in randomized controlled trials or observational studies'	Smith, S. C., Lamping, D. L., Banerjee, S., Harwood, R. H., Foley, B., Smith, P., et al. (2007). Development of a new measure of health-related quality of life for people with dementia: DEMQOL. Psychological Medicine, 37(5), 737–746	29 questions in tool DEMQOL score is total of 28 items Score range: 28 to 112. Reverse scoring on 5 items to give total score Recommend not to use in those with severe dementia = < 10 on MMSE (DEMQOL only) MMSE <17 is considered severe but for DEMQOL it is <10	higher scores indicate better HRQL	imputed missing data for respondents who had at least 50% of the remaining items complete using a person-specific mean using a widely accepted and established method of imputation (Ware et al. 1993, 1994) Ware, J. E., Kosinski, M. A. & Keller, S. D. (1994). SF-36 Physical and Mental Component Summary Measures : A User's Manual. The Health Institute, New England Medical Center: Boston. Ware, J. E., Snow, K. K., Kosinski, M. & Gandek, B. (1993). SF-36 Manual and Interpretation Guide. The Health Institute, New England Medical Center: Boston
DEMQOL Proxy	Smith, S. C., Lamping, D. L., Banerjee, S., Harwood, R. H., Foley, B., Smith, P., et al. (2007). Development of a new measure of health-related quality of life for people with dementia: DEMQOL. Psychological Medicine, 37(5), 737–746	31 items: Score range 31 to 124; higher overall total scores reflect better HRQL. Recommend use in mild, moderate and severe dementia	higher scores indicate better HRQL	imputed missing data for respondents who had at least 50% of the remaining items complete using a person-specific mean

MTSQ	Kristine M.C. Talley; Jean F. Wyman, Becky G. Olson-Kellogg; Ulf G. Bronas, Teresa C. McCarthy, Reliability and Validity of Two Measures of Toileting Skills in Frail Older Women Without Dementia. Journal of Gerontological Nursing• Vol. 42, No. 9, 2016	0-20 (max 5 questions **Response options: 0 =** none, **1 =** a little, **2=** some**, 3=** quite a lot**, and 4 =** cannot do	Higher scores = most difficulty. No MCID	No advice on missing data as this is new tool. In line with DEMQOL and other tools the Project Team decided to impute missing data for respondents who had at least 50% of the remaining items complete using a person-specific mean
Barthel Index	Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. *Int Disabil Stud*. 1988;10(2):61-63. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J*. 1965;14:61-65. Wade DT, Collin C. The Barthel ADL Index: a standard measure of physical disability? *Int Disabil Stud*. 1988;10(2):64-67.	0-20 Lower scores = greater disability. Scores >12 = supported independence Scores<8 = dependence	Higher scores=less disability	Could not find any guidance. However, since there is likely to be low rates of missing data, the Project Team decided to impute missing data for respondents who had at least 80% of the remaining items complete using a person-specific mean
MMSE	Folstein MF, Folstein SE, McHugh PR. Mini mental state. J Psychiatric Res 1975;12:189–98. Structural validity Rubright, J. D., Nandakumar, R., & Karlawish, J. (2015). Identifying an appropriate measurement modeling approach for the mini-mental state examination. Psychological Assessment, 28(2), 125–133. Blake, H., McKinney, M., Treece, K., Lee, E., & Lincoln, N. B. (2002). An evaluation of	Score 0-30, <23 = dementia scores between 0 and 9 indicating severe impairment, scores between 10 and 20 indicating moderate impairment and scores between 21 and 30 indicating mild	Higher scores = better performance and less cognitive impairment.	The suggested approach is to conduct item level MI but if not feasible then to use: • When most cases with missing data have <= 5 items missing, just use raw scores. • If many cases have more than 5 missing points, the goal of the analyses should also be considered: for descriptive analyses, use of scale-level MI including selected items; for regression analyses, raw scores can be used on their own or

	screening measures for cognitive impairment after stroke. Age and Ageing, 31(6), 451–456. Brayne, C. (1998). The mini-mental state examination, will we be using it in 2001? International Journal of Geriatric Psychiatry, 13(5), 285–290. Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. Journal of the American Geriatrics Society, 40(9), 922–35. Godin et al (2016) Handling missing Mini-Mental State Examination (MMSE) values: J Epi 27 (2017) 163-171	impairment.50 structural validity [23], predictive validity and reliability		in conjunction with scale-level MI. • A paper by Huppert 2005 seems to support this approach to small amounts of missing data. Depending on how much missingness there is overall and within a person we will select the most appropriate of these methods outline above