Support and treatment after joint replacement (STAR): Long-term follow-up of a care pathway for patients with long-term pain after knee replacement

Statistical Analysis Plan

Version 1.0 (21 December 2022)

The following people have reviewed the Statistical Analysis Plan and are in agreement with the contents				
Name	**Role**	**Title**	**Signature**	**Date**
Beverly A. Shirkey, PhD	Author	Research Associate in Medical Statistics	Beverly A. Shirkey, PhD	21 Dec 2022
Prof Tim Peters	Statistical Reviewer	Professor of Primary Care Health Services Research	T. J. Peters	21 Dec 2022
Prof Rachael Gooberman-Hill	Chief Investigator	Professor of Health and Anthropology	R. Gooberman-Hill	21 Dec 2022

Table of Contents

Abbreviations

Acronym	Detail
A&E	*Accident & emergency*
AE	*Adverse event*
CI	*Confidence interval*
DMC	*Data monitoring committee*
EQ-5D-5L	*EuroQol 5-dimension 5-level*
IQR	*Inter-quartile range*
ITT	*Intention to treat*
NHS	*National Health Service*
RCT	*Randomised controlled trial*
SAE	*Serious adverse event (subset of AE)*
SAP	*Statistical analysis plan*
SD	*Standard deviation*
SF-12	*Short form 12*
OKS	*Oxford Knee Score*
BPI	*Brief Pain Inventory*
DN-4	*Douleur Neuropathique 4*
HADS	*Hospital anxiety and depression scale*
CACE	*Complier average causal effect*
CRF	*Case report form*
CWP	*Chronic widespread pain*
CWP(M)	*Manchester's definition of CWP*
mice	*Multiple imputation by chained equation*

1. INTRODUCTION & PURPOSE

This document details the rules proposed and the presentation that will be followed, as closely as possible, when analysing and reporting the main results from the STAR Follow Up Trial.

The purpose of the plan is to:

1. Ensure that the analysis is appropriate for the aims of the trial, reflects good statistical practice, and that interpretation of a priori and post hoc analyses respectively is appropriate.
2. Explain in detail how the data will be handled and analysed to enable others to perform the actual analysis in the event of sickness or other absence.
3. Protect the project by helping it keep to timelines and within scope.

Additional exploratory or auxiliary analyses of data not specified in the protocol are permitted but fall outside the scope of this analysis plan (although such analyses would be expected to follow Good Statistical Practice).

The analysis strategy will be made available if required by journal editors or referees when the main papers are submitted for publication. Additional analyses suggested by reviewers or editors will, if considered appropriate, be performed in accordance with the Analysis Plan, but if reported the source of such a post-hoc analysis will be declared.

Editorial changes

Amendments to the statistical analysis plan will be described and justified in the final report of the trial in **Table 8** of this document.

Tables and figures

Throughout this document references are made to any skeleton tables and figures to be used in the reporting of the trial (e.g. **Figure F1** or **Table T1**). Such tables and figures can be found in **Appendix A** of this document, and are intended as a guide for trial reporting. Final versions of the tables/figures may differ: tables may be combined, and/or their layout or numbering may evolve. However the content will be consistent with **Appendix A.**

In this document, references to the follow up study protocol refer to "*0. STAR follow-up study Protocol v3 03-03-2021.docx" and the main study protocol refer to "*0. Protocol v9, 04-02-2019.pdf".

2. Trial Background and Objectives

2.1 Background

Please refer to the follow up study protocol, section titled "Background".

2.2. Trial objectives and aims

Please refer to the follow up study protocol, section titled "Aims and Objectives".

3. TRIAL DESIGN AND PROCEDURES

3.1. Trial design and configuration

Please refer to the main study protocol, section titled “Overview of trial design”, in the subsection titled “Main trial”.

3.2. Trial centres

Please refer to the follow up study protocol, page 1 “Study Sites”.

3.3. Selection criteria

Please refer to the follow up study protocol, section titled “Selection criteria”.

3.4. Description of interventions

Please refer to the main study protocol, section titled “Intervention: STAR Care Pathway”.

3.5. Control: Care as usual

Please refer to the main study protocol, section titled “Control: Usual Care”.

3.6. Randomisation procedures

Please refer to the main study protocol, section titled “Background”.

3.7. Sample size and justification

Please refer to the main study protocol, section titled “Sample size”.

3.8. Blinding and breaking of blind

Please refer to the main study protocol, section titled “Blinding”.

3.9. Trial committees

Please refer to the follow up study protocol, section titled “Trial organisation and oversight”.

3.10. Outcome measures

Please refer to the follow up study protocol, section titled “Outcome measures”.

4. GENERAL ANALYSIS CONSIDERATIONS

4.1. Analysis populations

All participants randomised to one or other arm of the trial will be analysed using the intention to treat (ITT) principal (in the presence of missing data this is strictly speaking referred to by the term 'as randomised') whereby participants are included in the group to which they were randomised, regardless of protocol deviations or non-compliance. Since these results run a risk of being biased if there are more than trivial amounts of missing outcome data, we will also perform (secondary) sensitivity analysis for the primary outcomes using the *rctmiss* command (White et al, 2017) in Stata to explore the effect of different means one arm at a time, best case/worst case analysis to show the extreme range of possible results, and *mice (Raghunathan et al, 2001; Van Buuren, 2007)* imputation.

4.2. Derived variables

Note, references to the primary and secondary outcomes are in the main and follow up study protocols. Also, note these are brief descriptions; details on handling missing data in terms of items required for the scores can be found in the missing data section and in the Stata code in the Appendix titled Stata code for derived variables.

Co-primary outcome measures

The Brief Pain Inventory is a questionnaire which consists of 14 questions, eleven of which are included in Section A of the follow-up questionnaire for STAR. The two scores that will be used as our co-primary are described below:

- **Pain Severity Score**: The Pain Severity Score is calculated by taking the mean of the rating scores of the first four questions in Section A of the questionnaire *(Q1+Q2+Q3+Q4)/4*.
- **Pain Interference Score**: The Pain Interference Score is calculated by taking the mean of the last seven questions in Section A of the questionnaire *(Q7+Q8+Q9+Q10+Q11+Q12+Q13)/7.*

Secondary outcome measures

- **Oxford Knee Score (OKS)**

The OKS is calculated using the items in section B of the questionnaire. To calculate the OKS we sum the responses to the 12 items (individual items scored 0-4, worst to best). The total score has a range of 0-48 (worst to best). The Oxford Knee Score can be split into two sub-scales: the pain and function subscales.

a. **OKS Pain subscale:** the raw subscale is equal to the sum of the responses to the following seven questions: 1, 4, 5, 6, 8, 9 and 10. This is then standardised to range from 0 to 100 by multiplying by 3.57 (100/(7*4)).

b. **OKS Function subscale:** the raw subscale is equal to the sum of the responses to the following five questions: 2, 3, 7, 11 and 12. This is then standardised to range from 0 to 100 by multiplying by 5 (100/(5*4)).

- **PainDETECT**

The PainDETECT score is calculated using items in Section C of the questionnaire. The first seven questions are scored zero to five (Never – Very Strong). The eighth question is a picture representation of the pain and these are scored between negative one and positive one. Lastly the ninth question is scored 2 if "Yes" is selected and zero if "No" is selected. The sum of each score provides the PainDETECT score. This ranges from -1 to 38 and scores fall into three categories: (-1 to 12) nociceptive, (13-18) unclear and (19-38) neuropathic pain.

- **DN-4** (Douleur Neuropathique)

The DN-4 score is a score out of 7 corresponding to the number of 'yes' answers the patient gave in Section C.

- **EQ-5D-5L (analyses of the EQ-5D-5L will be covered in the Heath Economics Analysis Plan)**

The EQ-5D-5L will be analyses as part of the economic evaluation and does not form part of the secondary outcome set for the purposes of this SAP.

- **Short Form-12**

This outcome is derived by software provided by **QualityMetric Incorporated, LLC** using responses from Section E of the questionnaire.

- **Hospital Anxiety and Depression Scale (HADS)**

HADS is split into two sub-scales, the Anxiety Scale and the Depression Scale. Each scale comprises of the sum of responses from 7 items from Section G of the questionnaire. Each item is scored from 0 to 3 with 0 being the best case scenario and 3 being the worst. Each of the two sub-scales are categorised into a normal score (0-7); borderline anxiety/depression (8-10) and clinical anxiety/depression (≥11).

- **ICECAP-A**

ICECAP-A uses responses from Section H of the questionnaire and provides a state of 5 characteristics 'XXXXX'. This then allows us to calculate a tariff value for items which make up the state. This tariff value is the sum of pre-specified values corresponding to the answers given in the questionnaire (Flynn et al, 2015). The code for this is presented in the appendix.

- **Pain Catastrophizing Scale**

The Pain Catastrophizing Scale is split into three sub-scales, The Rumination Scale, The Magnification Scale and The Helplessness Scale. Each scale is a sum of the ratings given to each of the following items of Section I of the questionnaire:

1. The Rumination Scale: 8, 9, 10, 11
2. The Magnification Scale: 6, 7, 13
3. The Helplessness Scale: 1, 2, 3, 4, 5, 12

The whole scale is additive of the three subscales and will be used for the secondary outcomes analysis.

- **Pain Solutions Questionnaire (PaSol)**

The Pain Solution Questionnaire is split into four sub-scales – Solving Pain, Meaningfulness of Life despite Pain, Acceptance of Insolubility of Pain and Belief in Solution. Each scale is a sum of the answers given to each of the following items of Section J of the questionnaire:

1. Solving Pain: 7, 10, 11, 12
2. Meaningfulness of Life despite Pain: 1, 2, 3, 8, 13
3. Acceptance of Insolubility of Pain: 4, 5, 9
4. Belief in Solution: 6, 14

The four sub-scales will be analysed separately.

- **Self-Administered Patient Satisfaction Scale for Primary Hip and Knee Arthroplasty**

The satisfaction scale is made up of four questions focusing on satisfaction with the extent of pain relief, improvement in ability to perform home or yard work, ability to perform recreational activities, and overall satisfaction with joint replacement. Items are scored on a 4-point Likert scale with response categories consisting of very satisfied (100 points), somewhat satisfied (75 points), somewhat dissatisfied (50 points), and very dissatisfied (25 points). The scale is calculated by taking an unweighted average over these four questions providing a score ranging from 25 to 100 (with 100 being most satisfied). This will be treated as a continuous variable in the analysis.

- **Body Map**

The body map in Section M of the questionnaire is used to determine chronic widespread pain according to Manchester's definition CWP(M). Patients indicate sections of the body where they feel pain by shading in sections of a mannequin (viewed from front, back, left

and right) and the Manchester definition is used to categorise patients into those who have CWP(M) and those who do not. To satisfy the Manchester definition of chronic widespread pain (CWP(M)), pain must be reported in at least two sections of each two contralateral limbs and in the axial skeleton and have been present for at least 3 months. Although the presence of pain for 3 month is not recorded in the trial, we will classify patients based on the other elements of the definition.

4.3. Procedures for missing data

In all tables missing data will be indicated. If the amount of missing data differs substantially between treatment groups potential reasons will be explored. Sensitivity analyses will be conducted (including through the use of multiple imputation by chained equations (mice) methods) to examine the influence of missing data on the key trial findings. When using mice, 25 datasets will be generated and 10 switching procedures undertaken. The imputation model will include all variables predictive of missingness, together with all of the variables included in the main substantive model. Comparisons of results from 'as randomised' analyses of complete cases with 'as randomised' analyses where missing data were imputed will be presented in **Table 5 and Table 6**

The model used for imputation will include a baseline measure of the outcome, any other observations of the outcome at different follow-up times, randomisation group, age/gender, centre and any other restriction variables for the randomisation (i.e. stratification/minimisation), we will consider also including any other variables that are either strongly associated with missingness or likely to have some prognostic value. This list will be finalised before conducting the mice analyses.

BPI (severity and interference): The first four items of section A must be complete to calculate the score for the severity scale. Four of the last seven items of section A of the questionnaire must be complete to calculate the interference scale by averaging complete items.

OKS: If 1 or 2 questions are missing, then the mean value can be used to fill the gaps.

PainDETECT: If any of the first seven items of section C are missing impute with the mean of the complete items in the first seven items. If question 8 of section C is missing do not add or subtract anything from the score (i.e. treat the value of that item as zero). If question 9 is missing, assume the response is no, thus, treat the value of the item as zero.

DN-4: No score can be calculated if more than 4 items are missing. The score is a proportion of "Yes" responses.

Short form-12: The short form-12 requires 50% of items to be completed.

HADS: The score for a single missing item from a sub-scale is inferred by using the mean of the remaining six items. If more than one item is missing from a sub-scale, that sub-scale cannot be calculated.

ICECAP-A: There is not any internal way of dealing with missing data, as each attribute on the questionnaire is intended to be mutually exclusive.

Pain Catastrophizing Scale: There are no formal guidelines for dealing with missing data in the PCS. We allow one item to be missing from each subscale and this item will be replaced by the mean of the complete items in that subscale.

Pain Solution Score: 75% of items in each subscale need to be complete in order to calculate a score. We extrapolate the score to new total sub-scores. For example, if 4 items of 5 have been completed. The total score of the 4 is divided be 4 and multiplied by 5.

Satisfaction scale: There are no formal guidelines for dealing with missing data. If one item is missing we will fill the missing value with the mean of the completed items. If more than one item is missing we will not calculate the score.

DESCRIPTION OF PARTICIPANT CHARACTERISTICS

4.1. Disposition

A flow of patients through the trial will be summarised in a CONSORT diagram **(Figure 1 and Figure 2**) that will include the eligibility, exclusion, number of patients randomised to the two treatment groups, loss to follow up and the number of patients analysed.

4.2. Baseline characteristics

Baseline characteristics have been reported in the main trial.

5. ASSESSMENT OF TRIAL QUALITY

5.1. Eligibility checks

Eligibility checks were reported in the main trial.

5.2. Data validation

The system will incorporate data entry and validation rules to reduce data entry errors, and management functions to facilitate auditing and data quality assurance. We will use a secure, web-based data collection platform (REDCap) which will be developed, validated, hosted and supported by the University of Bristol.

Named blinded assessors perform data completeness checks of data and contact patients if there is missing data in their questionnaires. This will reduce the amount of missing data as patients will have the opportunity to complete missing items over the phone. This may also be an opportunity to clarify any misunderstandings in the questionnaire. It is important for these telephone calls to be done by a member of a different trial centre team so that the researcher who phones the patient is unaware of the treatment group allocation, to minimise bias.

All data are received centrally directly from the participant in the form of self-completion questionnaires or collected over the phone by blinded assessors. Data are entered into REDCap by members of the co-ordinating centre research team.

5.3. Trial completion

Withdrawals are summarised in in Table 1.

5.4. Compliance

There was no intervention in the follow up phase, and compliance in the main trial has been analysed in the main trial report.

5.7. Protocol deviations

These have already been fully covered in the original analyses and reporting of the main trial.

5.8. Specify and justify changes made to the planned statistical analyses

Any adjustment to the statistical analysis plan will be logged in **Table 7.**

6. ANALYSIS OF EFFECTIVENESS

6.1 Statistical analysis

STATA 17.0 will be used for all statistical analyses.

Continuous variables will be summarised using the mean (arithmetic or if necessary geometric) and standard deviation (SD), and the median and inter-quartile range (IQR), and categorical data will be summarised as a number and percentage.

6.2. Summary of primary and secondary outcomes

All primary and secondary analyses will be conducted on an 'as randomised' basis, using the appropriate regression model. Assumptions for each regression will be checked to make sure the correct method of analysis is being used. A summary of the primary and secondary outcomes can be seen in **Table 2.**

6.3. Primary analysis

Each of the co-primary outcome measures, BPI Severity and BPI Interference scales, will be analysed to compare treatment groups using linear regression. The models will be adjusted for trial centre and baseline pain scores. Estimates will be calculated of the effect that intervention has on each of the BPI scores compared with usual care.

Results from the primary analysis will be presented in **Table 3**.

6.4. Secondary analyses

The secondary outcomes will be analysed using appropriate regression models in a similar manner to the primary analysis. A summary of the primary and secondary outcomes can be seen in **Table 17.**

Results from the secondary analysis will be presented in **Table 4**.

6.5. Sensitivity analysis

We will investigate the influence of missing data using sensitivity analyses that make different assumptions, such as best/worst case scenarios and rctmiss graphics, as well as using multiple imputation by chained equation (mice) to impute missing data. We will also explore using reasons given for missing data, when available, to impute plausible values.

6.6. Subgroup analyses

We recognise that there will be low power for subgroup analyses and therefore only cautious indications of potential need for further research will be drawn from them. Subgroup analyses will be performed by introducing appropriate interaction terms in the regression models, to investigate any differential effects according to the same pre-defined factors as considered before – namely, trial centre and continuous versions of the Oxford Knee Score and Pain Solutions Questionnaire at baseline.

7. ANALYSIS OF SAFETY

7.1. Adverse reactions

No safety data were collected for the long-term follow up study. However, the number of deaths reported in the vital status check before sending out the follow up questionnaires will be reported, plus the number and percentage per arm. The total deaths from the main study and follow-up study will be combined per arm and reported. As death was not an outcome of this study, no formal statistical test will be performed.

8. FINAL REPORT TABLES AND FIGURES

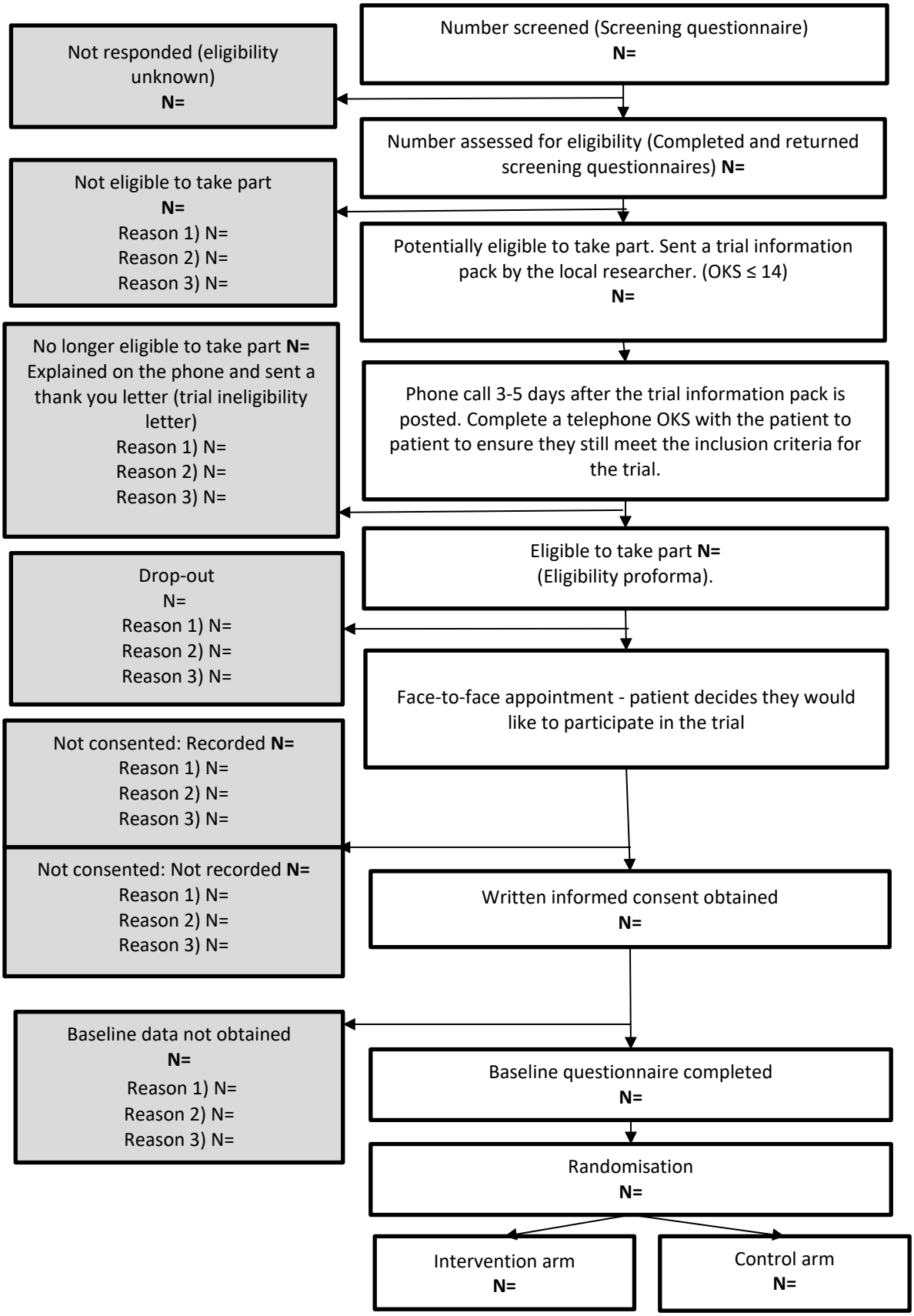

Figure 1: Consort flow diagram to monitor the number or patients included in the trial up to randomisation

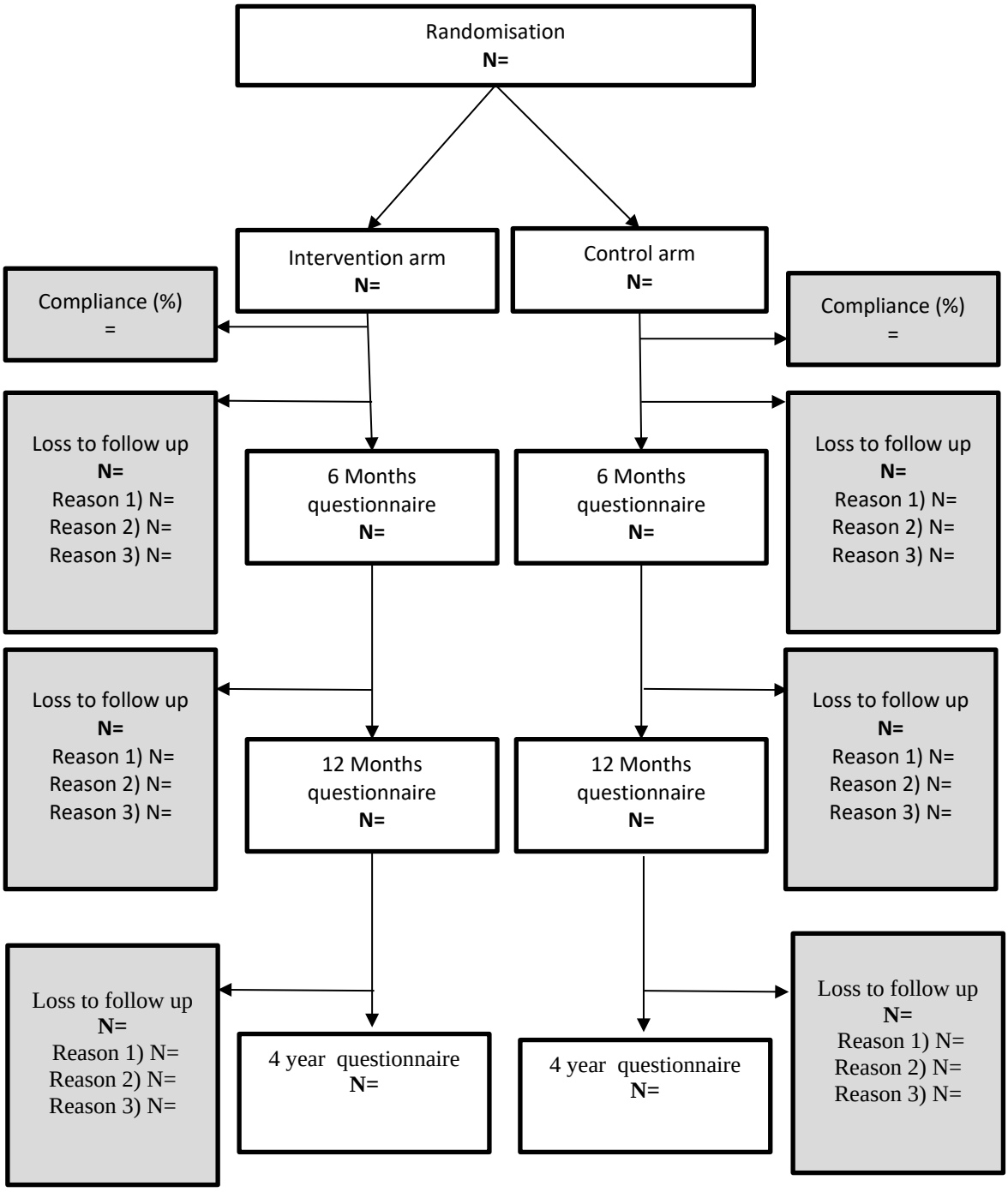

Figure 2: Consort diagram to monitor the number of patients included in the trial post randomisation

Tables

9.1 Follow up response

Table 2: Reasons given for not responding to 4 year follow up by arm

Reason	N (%)	Usual Care (%)	Intervention (%)
Withdrawn from main study			
In too much pain			
In no pain			
Other reason 1			
Other reason 2….			
….			
No reason given			
Total			

9.2. Outcome summaries

Table 2: Primary and Secondary endpoint summary

Outcome measure	Type of data	Range of values
BPI – Pain severity scale	Continuous	0-10 (best to worst)
BPI – Pain Interference scale	Continuous	0-10 (best to worst)
Oxford Knee Score (OKS)	Continuous	0-48 (worst to best)
Oxford Knee Score (OKS) Pain Subscale (7 items)	Continuous	1-100 (worst to best)
Oxford Knee Score (OKS) Function Subscale (7 items)	Continuous	1-100 (worst to best)
Douleur Neuropathique 4 (DN-4)	Continuous	0-7 (best to worst)
PainDETECT	Continuous	-1-38 (best to worst)
Hospital Anxiety and Depression Scale (HADS)	Continuous	Each subscale: 0-21 (best to worst)
Hospital Anxiety and Depression Scale (HADS)	Ordinal	normal score (0-7); borderline anxiety/depression (8-10) and clinical anxiety/depression (≥11)
Pain Catastrophizing Scale	Continuous	The Rumination Scale: 0-16 (best to worst) The Magnification Scale: 0-12 (best to worst) The Helplessness Scale: 0-24 (best to worst) Whole score: 0-52 (best to worst)
Pain Solution Questionnaire (PaSol)	Continuous	Solving Pain: 0-24 (worst to best) Meaningfulness of Life despite Pain: 0-30 (worst to best) Acceptance of Insolubility of Pain: 0-18 (worst to best) Belief in Solution: 0-12 (worst to best)
Self-Administered Patient Satisfaction Scale for Primary Hip and Knee Arthroplasty	Continuous	25-100 (worst to best) Items are scored on a 4-point Likert scale with response categories consisting of very satisfied (100 points), somewhat satisfied (75 points), somewhat dissatisfied (50 points), and very dissatisfied (25 points).
ICECAP-A	Continuous	-0.001 to 1 (worst to best)
Short Form-12	Continuous	
Body Map	Binary	0/1: CWP(M) or not
Q5 Section A	Ordinal	“Rarely”, “Sometimes”, “Often”, “Most of the time”, “All of the time”
Q8 Section D	Ordinal	“Rarely”, “Sometimes”, “Often”, “Most of the time”, “All of the time”
Q5 Section L	Ordinal	“Much better”, “A bit better”, “The same”, “A bit worse”, “Much worse”
Resource use (reported by HE)		
EQ-5D-5L (reported by HE)		

9.3. Primary outcome results

Table 3: Primary outcome table

	N	Mean at Baseline, 4 year	SD at Baseline, 4 year	Difference in means[1]	95% CI	P-value
BPI Severity						
BPI Interference						

[1] Adjusted for trial centre and baseline Brief Pain Inventory Severity and Interference Scales

9.4. Secondary outcomes results

Table 4: Secondary outcomes tables

	N	Mean	SD	Difference in means[1]	95% CI	P-value
BPI Severity						
BPI Interference						
OKS						
DN-4						
PainDETECT						
Pain Catastrophizing scale						
PaSol: Solving Pain						
PaSol: Meaningful life						
PaSol: Acceptance of pain						
PaSol: Belief in solution						
Patient Satisfaction						
ICECAP-A						
Short form-12						
HADS: Anxiety						
HADS: Depression						
Section A: Question 5						
Section D: Question 8						
Section L: Question 5						
	N	**Odds ratio[1]**	**95% CI**	**P-value**		
Body Map (CWP(M))						

[1] Adjusted for trial centre and baseline Brief Pain Inventory Severity and Interference Scales

9.5. Sensitivity analysis for primary endpoint

Comparison of results of ITT analysis of complete cases with ITT analysis where missing data were imputed using “best” and “worst” case scenarios and the method of mice for primary outcome of **BPI Severity Score**.

Table 5:Sensitivity analysis for missing data in BPI Severity Score.

	N	Difference in means[a]	95% CI	p-value
Complete case				
Best case / worst favouring Intervention				
Best case / worst case favouring control				
mice				
Using imputed values for those reporting pain or fine with plausible values and mice for those not reporting reasons				

[a] *Adjusted for trial centre and for baseline OKS*

Comparison of results of ITT analysis of complete cases with ITT analysis where missing data were imputed using “best” and “worst” case scenarios and the method of mice for primary outcome of **BPI Interference Score**.

Table 63: Sensitivity analysis for missing data BPI Interference Score.

	N	Difference in means[a]	95% CI	p-value
Complete case				
Best case / worst favouring Intervention				
Best case / worst case favouring control				
mice				
Using imputed values for those reporting pain or fine with plausible values and mice for those not reporting reasons				

[a] *Adjusted for trial centre and for baseline OKS*

Figure 3: **RCTmiss graphic (similar to example below)**

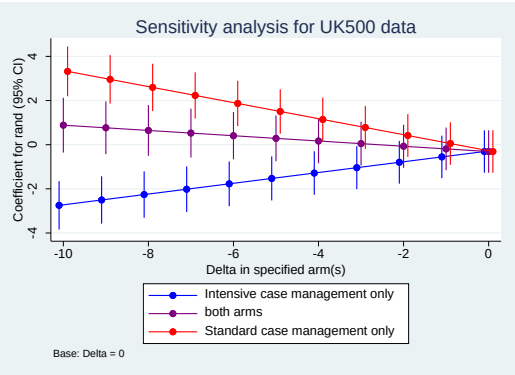

Figure 4: RCTmiss graphic

Amendments to the SAP

Table 7: Amendments to the SAP

Previous version	Previous date	New version	New date	Brief summary of changes

9. APPENDICES

9.1. Stata code for derived variables

```
*BPI_severity

 * (note: requires all 4 items))
* count non-missing
egen bpi_sev_non_miss_ltfu = ///
 rownonmiss(worst_ltfu least_ltfu average_ltfu rightnow_ltfu)
gen bpi_severity_ltfu = ///
 (worst_ltfu + least_ltfu + average_ltfu + rightnow_ltfu)/4 ///
 if bpi_sev_non_miss_ltfu==4


*BPI_interference (allow for up to 3 missing items)
* count non-missing
egen bpi_int_non_miss_ltfu = ///
 rownonmiss(interfere_gen_ltfu interfere_mood_ltfu interfere_walk_ltfu  ///
            interfere_norm_ltfu interfere_relation_ltfu interfere_sleep_ltfu ///
                      interfere_life_ltfu)

egen  bpi_int_ltfu = rowmean(
 interfere_gen_ltfu interfere_mood_ltfu interfere_walk_ltfu + ///
  interfere_norm_ltfu  interfere_relation_ltfu  interfere_sleep_ltfu ///
  interfere_life_ltfu) if range(bpi_int_non_miss_ltfu, 4, 7)


*Oxford knee score
* note - not sure why replaced was used in name instead of oks -
* these are the entered oks answers - coded 0-4
* OKS: If 1 or 2 questions are missing, then the mean value can be used to fill the gaps
egen oks_non_miss_ltfu = rownonmiss( ///
replaced_pain_ltfu replaced_wash_ltfu replaced_car_ltfu ///
replaced_walk_ltfu replaced_sat_ltfu replaced_limp_ltfu ///
replaced_kneel_ltfu replaced_trouble_ltfu replaced_work_ltfu ///
replaced_giveway_ltfu replaced_shop_ltfu replaced_stairs_ltfu)


egen oks_ltfu_average = rowmean( ///
replaced_pain_ltfu replaced_wash_ltfu replaced_car_ltfu ///
replaced_walk_ltfu replaced_sat_ltfu replaced_limp_ltfu ///
replaced_kneel_ltfu replaced_trouble_ltfu replaced_work_ltfu ///
replaced_giveway_ltfu replaced_shop_ltfu replaced_stairs_ltfu) ///
if inrange(oks_non_miss_ltfu, 10, 12)

generate oks_ltfu = round(12*oks_ltfu_average, 1)

* OKS pain subscale
* again - assume if 1 or 2 questions missing, then use mean value to fill gaps
egen oks_pain_non_miss_ltfu = rownonmiss( ///
replaced_pain_ltfu replaced_walk_ltfu replaced_sat_ltfu replaced_limp_ltfu ///
replaced_trouble_ltfu replaced_work_ltfu replaced_giveway_ltfu)

egen oks_pain_ltfu_average = rowmean( ///
replaced_pain_ltfu replaced_walk_ltfu replaced_sat_ltfu replaced_limp_ltfu ///
replaced_trouble_ltfu replaced_work_ltfu replaced_giveway_ltfu) ///
if inrange(oks_pain_non_miss_ltfu, 5,7)

* generate sum score and scale to 100
generate oks_pain_ltfu = round(7*oks_pain_ltfu_average, 1)*(100/(7*4))


*OKS function subscale
egen oks_func_non_miss_ltfu = rownonmiss( ///
 replaced_wash_ltfu replaced_car_ltfu replaced_kneel_ltfu  ///
 replaced_shop_ltfu replaced_stairs_ltfu)

egen oks_func_ltfu_average = rowmean( ///
 replaced_wash_ltfu replaced_car_ltfu replaced_kneel_ltfu  ///
 replaced_shop_ltfu replaced_stairs_ltfu) ///
if inrange(oks_func_non_miss_ltfu, 3,5)


* generate sum score and scale to 100
generate oks_func_ltfu = round(5*oks_func_ltfu_average, 1)*(100/(5*4))
```

```
*DN-4

egen dn_4_non_miss_ltfu = rownonmiss( ///
       feelpain_burn_ltfu feelpain_cold_ltfu feelpain_shocks_ltfu ///
       feelpain_tingling_ltfu feelpain_pinsneedles_ltfu feelpain_numb_ltfu ///
       feelpain_itch_ltfu)

egen dn_4_ltfu_total = rowtotal( ///
       feelpain_burn_ltfu feelpain_cold_ltfu feelpain_shocks_ltfu ///
       feelpain_tingling_ltfu feelpain_pinsneedles_ltfu feelpain_numb_ltfu ///
       feelpain_itch_ltfu)  ///
if inrange(dn_4_non_miss_ltfu,3,7)

generate dn_4_ltfu = dn_4_ltfu_total/dn_4_non_miss_ltfu  ///
if inrange(dn_4_non_miss_ltfu,3,7)


*PainDETECT

* 7 questions (scored 1-5)
/*PainDETECT: If any of the first seven items of section C are
 missing impute with the mean of the complete items in the first seven items.
 If question 8 of section C is missing do not add or subtract anything
 from the score (i.e. treat the value of that item as zero).
 If question 9 is missing, assume the response is no, thus,
 treat the value of the item as zero.
 */

egen pd_first_7_non_miss_ltfu = rownonmiss( ///
painfeel_burning_ltfu painfeel_ting_ltfu painfeel_touch_ltfu ///
painfeel_sudden_ltfu painfeel_temp_ltfu painfeel_sensation_ltfu ///
painfeel_pressure_ltfu)

egen pd_first_7_ltfu_average = rowmean( ///
painfeel_burning_ltfu painfeel_ting_ltfu painfeel_touch_ltfu ///
painfeel_sudden_ltfu painfeel_temp_ltfu painfeel_sensation_ltfu ///
painfeel_pressure_ltfu) ///
if inrange(pd_first_7_non_miss_ltfu,1,7)


tab painfeel_describe_ltfu painfeel_raidate_ltfu if pd_first_7_non_miss_ltfu==0

* multiply by 7 and round
generate pd_first_7_ltfu = round(7*pd_first_7_ltfu_average, 1)

* score picture (question 8)
/*
Please select the picture that best describes the course of your pain:
 Persistent pain with slight fluctuations 0
 Persistent pain with pain attacks -1
 Pain attacks without pain between them +1
 Pain attacks with pain between them +1
 */
codebook painfeel_describe_ltfu
 recode painfeel_describe_ltfu (1=0) (2 = -1) (3=1) (4=1),  ///
 generate(pd_course_ltfu)
 bysort painfeel_describe_ltfu: tab pd_course_ltfu

 * add in question 8 and 9(painfeel_raidate_ltfu) - treat as zero as missing
 * from stata help
 * note egen, missing...  It creates the (row) sum of the variables in varlist,
 * treating missing as 0.  If missing option is specified and all values in
 * varlist are missing for an observation, newvar is set to missing
 * for that observation.

 egen pain_detect_ltfu  = ///
  rowtotal(pd_first_7_ltfu pd_course_ltfu painfeel_raidate_ltfu) ///
  , missing

 * check to make sure at least one of the 7 questions is answered
 assert  pain_detect_ltfu ==. if pd_first_7_non_miss_ltfu==0
```

```
*HADS

*HADS: The score for a single missing item from a sub-scale is inferred
* by using the mean of the remaining six items. If more than one item is
*  missing from a sub-scale, that sub-scale cannot be calculated.

egen hads_anx_non_miss_ltfu = rownonmiss( ///
       mood_wound_ltfu mood_fright_ltfu mood_worry_ltfu mood_relax_ltfu ///
       mood_butterfly_ltfu mood_restless_ltfu mood_panic_ltfu)

egen hads_anx_ltfu_ave = rowmean( ///
       mood_wound_ltfu mood_fright_ltfu mood_worry_ltfu mood_relax_ltfu ///
       mood_butterfly_ltfu mood_restless_ltfu mood_panic_ltfu)  ///
if inrange(hads_anx_non_miss_ltfu,6,7)

generate hads_anx_ltfu = round(7*hads_anx_ltfu_ave, 1)

egen hads_depr_non_miss_ltfu = rownonmiss( ///
       mood_wound_ltfu mood_fright_ltfu mood_worry_ltfu mood_relax_ltfu ///
       mood_butterfly_ltfu mood_restless_ltfu mood_panic_ltfu)

egen hads_depr_ltfu_ave = rowmean( ///
       mood_wound_ltfu mood_fright_ltfu mood_worry_ltfu mood_relax_ltfu ///
       mood_butterfly_ltfu mood_restless_ltfu mood_panic_ltfu)  ///
if inrange(hads_depr_non_miss_ltfu,6,7)

generate hads_depr_ltfu = round(7*hads_depr_ltfu_ave,

*Pain Catastrophizing Scale
* only one item missing in subscale -
* replace missing item with average of those left

 * allow for only 1 item missing in subscale,
 * average other items to impute one missing item

 global pcs_r_items = ///
 "pain_away_ltfu pain_mind_ltfu pain_hurts_ltfu pain_stop_ltfu"
 global pcs_m_items = ///
 "pain_worse_ltfu pain_events_ltfu pain_serious_ltfu"
global pcs_h_items = ///
 "pain_worry_ltfu pain_can_go_on_ltfu pain_terrible_ltfu pain_awful_ltfu pain_stand_more_ltfu
pain_intensity_ltfu"


foreach ss in r m h {
 local items =         "pcs_`ss'_items"
 di "$`items'"
 egen pcs_`ss'_non_miss_ltfu = rownonmiss($`items')
 tab pcs_`ss'_non_miss_ltfu
local item_num: word count $`items'
di `item_num'
egen pcs_`ss'_sum =  rowtotal($`items')
generate psc_`ss'_ltfu = cond(pcs_`ss'_non_miss_ltfu==`item_num', ///
       pcs_`ss'_sum, .)
replace psc_`ss'_ltfu  = pcs_`ss'_sum * (1 + (1/(`item_num'-1))) ///
if `item_num'- pcs_`ss'_non_miss_ltfu==1,
list $`items' pcs_`ss'_sum psc_`ss'_ltfu ///
 if `item_num'- pcs_`ss'_non_miss_ltfu==1 ///
, noobs nolab abbr(32)
 }

 * require all 3 subscales
gen pcs_total_ltfu = psc_r_ltfu + psc_m_ltfu + psc_h_ltfu ///
 if psc_r_ltfu<. & psc_m_ltfu<. & psc_h_ltfu<.
```

```
 *PaSol
 * (subscale requires 75% completed items)
  * average other items to impute one missing item

 * two subscales can have one missing item
 global pa_sol_solve_items = ///
 " deal_pain_search_ltfu deal_pain_rid_ltfu deal_pain_solut_ltfu deal_pain_without_ltfu "
 global pa_sol_meaning_items = ///
 "deal_pain_meaningful_ltfu deal_pain_wayout_ltfu deal_pain_live_ltfu deal_pain_best_ltfu
deal_pain_way_ltfu"

foreach ss in solve meaning {
 local items =         "pa_sol_`ss'_items"
 di "$`items'"
 egen pa_sol_`ss'_non_miss_ltfu = rownonmiss($`items')
 tab pa_sol_`ss'_non_miss_ltfu
local item_num: word count $`items'
di `item_num'
egen pa_sol_`ss'_sum =  rowtotal($`items')
generate pa_sol_`ss'_ltfu = cond(pa_sol_`ss'_non_miss_ltfu==`item_num', ///
        pa_sol_`ss'_sum, .)
replace pa_sol_`ss'_ltfu  = pa_sol_`ss'_sum * (1 + (1/(`item_num'-1))) ///
if `item_num'- pa_sol_`ss'_non_miss_ltfu==1,
list $`items' pa_sol_`ss'_sum pa_sol_`ss'_ltfu ///
 if `item_num'- pa_sol_`ss'_non_miss_ltfu==1 ///
, noobs nolab abbr(32)
 }

* requires all
gen pa_sol_accept_ltfu =  ///
deal_pain_no_solution_ltfu + deal_pain_cntrl_ltfu + deal_pain_accept_ltfu
* requires all
gen pa_sol_belief_ltfu = deal_pain_conf_ltfu + deal_pain_treat_ltfu

*ICECAP-A
(required all 5 items to be completed)

matrix UTILS=(-0.001,0.101,0.191,0.222\/*
*/-0.024,0.096,0.189,0.228\/*
*/0.006, 0.084, 0.156, 0.188\/*
*/0.021, 0.091, 0.159, 0.181\ /*
*/ -0.003, 0.069, 0.154, 0.181)
gen sta_index=UTILS[1,feel_settled_bl[_n]]
gen att_index=UTILS[2,feel_love_bl[_n]]
gen aut_index=UTILS[3,mood_indep_bl[_n]]
gen ach_index=UTILS[4,mood_achieve_bl[_n]]
gen enj_index=UTILS[5,mood_pleasure_bl[_n]]
gen tariff=sta_index+att_index+aut_index+ach_index+enj_index


*Satisfaction scale
* 1 item can be missing

 egen satis_scale_non_miss_ltfu = rownonmiss( ///
 satisfied_surgery_ltfu satis_improve_ltfu satis_housework_ltfu satis_leisure_ltfu)
 tab satis_scale_non_miss_ltfu
local item_num: word count ///
satisfied_surgery_ltfu satis_improve_ltfu satis_housework_ltfu satis_leisure_ltfu
di `item_num'
egen satis_scale_sum =  rowtotal ///
(satisfied_surgery_ltfu satis_improve_ltfu satis_housework_ltfu satis_leisure_ltfu)
generate satis_scale_ltfu = cond(satis_scale_non_miss_ltfu ==`item_num', ///
        satis_scale_sum, .)
replace satis_scale_ltfu   = satis_scale_sum * (1 + (1/(`item_num'-1))) ///
if `item_num'- satis_scale_non_miss_ltfu==1,
list ///
satisfied_surgery_ltfu satis_improve_ltfu satis_housework_ltfu ///
 satis_leisure_ltfu ///
satis_scale_sum satis_scale_ltfu  ///
 if `item_num'- satis_scale_non_miss_ltfu==1 ///
, noobs nolab abbr(32)

*EQ-5D-5L
* Calculated by health economics.

*ShortForm-12
* Calculated by software provided by QualityMetric Incorporated, LLC.
```

```
*body mapping tool
*CWP(M)
foreach x in 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29{
gen mannequin_BL_`x' = mannequin_bl___`x' if mannequin_bl___`x' != .
replace mannequin_BL_`x' = mannequin_bl_v2___`x' if mannequin_bl_v2___`x' != .
}
foreach x in 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29{
tab mannequin_BL_`x'
}


egen mannequin_count_BL = rowtotal(mannequin_BL_1 mannequin_BL_2 mannequin_BL_3 mannequin_BL_4
mannequin_BL_5 mannequin_BL_6 mannequin_BL_7 mannequin_BL_8 mannequin_BL_9 mannequin_BL_10
mannequin_BL_11 mannequin_BL_12 mannequin_BL_13 mannequin_BL_14 mannequin_BL_15
mannequin_BL_16 mannequin_BL_17 mannequin_BL_18 mannequin_BL_19 mannequin_BL_20
mannequin_BL_21 mannequin_BL_22 mannequin_BL_23 mannequin_BL_24 mannequin_BL_25
mannequin_BL_26 mannequin_BL_27 mannequin_BL_28 mannequin_BL_29)
tab mannequin_count_BL if redcap_event_name == "01_arm_1"
tab mannequin_count_BL
*101 patients reported zero painful regions
gen mannequin_count_zero = 1 if mannequin_count_BL == 0
tab mannequin_count_zero
replace mannequin_count_BL = . if mannequin_count_BL == 0



tab ou_side mannequin_BL_10
gen mannequin_pain_in_TKR_knee = 1 if ou_side == 1 & mannequin_BL_10 == 1
tab ou_side mannequin_BL_14
replace mannequin_pain_in_TKR_knee = 1 if ou_side == 2 & mannequin_BL_14 == 1
tab mannequin_pain_in_TKR_knee

replace mannequin_count_BL = mannequin_count_BL - 1 if ou_side == 1 & mannequin_BL_10 == 1
replace mannequin_count_BL = mannequin_count_BL - 1 if ou_side == 2 & mannequin_BL_14 == 1
tab mannequin_count_BL
gen mannequin_pain_in_TKR_knee_only = 1 if mannequin_count_BL == 0
tab mannequin_pain_in_TKR_knee_only

gen mannequin_count_BL_cat = 0 if mannequin_count_BL == 0
replace mannequin_count_BL_cat = 1 if mannequin_count_BL == 1
replace mannequin_count_BL_cat = 2 if mannequin_count_BL == 2
replace mannequin_count_BL_cat = 3 if mannequin_count_BL == 3
replace mannequin_count_BL_cat = 4 if mannequin_count_BL == 4
replace mannequin_count_BL_cat = 5 if mannequin_count_BL >= 5 & mannequin_count_BL != .
tab mannequin_count_BL_cat, mi


/*

mannequin_BL_1 - right shoulder
mannequin_BL_2 - right elbow
mannequin_BL_3 - right forearm
mannequin_BL_4 - right hand
mannequin_BL_5 - left shoulder
mannequin_BL_6 - left elbow
mannequin_BL_7 - left forearm
mannequin_BL_8 - left hand
mannequin_BL_9 - right thigh
mannequin_BL_10 - right knee
mannequin_BL_11 - right shin
mannequin_BL_12 - right foot
mannequin_BL_13 - left thigh
mannequin_BL_14 - left knee
mannequin_BL_15 - left shin
mannequin_BL_16 - left foot
mannequin_BL_17 - head
mannequin_BL_18 - throat/centre chest
mannequin_BL_19 - right breast/chest
mannequin_BL_20 - left breast/chest
mannequin_BL_21 - tummy
mannequin_BL_22 - left upper back
mannequin_BL_23 - right upper back
mannequin_BL_24 - centre upper back
mannequin_BL_25 - lower left back
mannequin_BL_26 - lower right back
mannequin_BL_27 - lower centre back
mannequin_BL_28 - left bum cheek
mannequin_BL_29 - right bum cheek
```

```
*/

*For subjects to satisfy the Manchester definition of chronic widespread pain [CWP(M)],
*pain must be reported in at least two sections of two contralateral limbs and in the axial
*skeleton, and have been present for at least 3 months [6]. This is in contrast to the ACR
definition
*of CWP which requires only that pain be present in any part of contralateral body quadrants,
in
*addition to axial pain.



egen axial_skeleton = rowtotal(mannequin_BL_25 mannequin_BL_26 mannequin_BL_27 mannequin_BL_24
mannequin_BL_18 mannequin_BL_19 mannequin_BL_20 mannequin_BL_22 mannequin_BL_23),  m
tab axial_skeleton, mi
egen Rleg = rowtotal(mannequin_BL_9 mannequin_BL_10 mannequin_BL_11 mannequin_BL_12
mannequin_BL_29),  m
tab Rleg, mi
egen Lleg = rowtotal(mannequin_BL_13 mannequin_BL_14 mannequin_BL_15 mannequin_BL_16
mannequin_BL_28),  m
tab Lleg, mi
egen Rarm = rowtotal(mannequin_BL_1 mannequin_BL_2 mannequin_BL_3 mannequin_BL_4),  m
tab Rarm, mi
egen Larm = rowtotal(mannequin_BL_5 mannequin_BL_6 mannequin_BL_7 mannequin_BL_8),  m
tab Larm, mi


gen Rarm_Lleg = 1 if Rarm >= 2 & Lleg >= 2 & Rarm != . & Lleg != .
replace Rarm_Lleg = 0 if Rarm < 2 & Lleg < 2
tab Rarm_Lleg, mi
gen Larm_Rleg = 1 if Larm >= 2 & Rleg >= 2  & Larm != . & Rleg != .
replace Larm_Rleg = 0 if Larm < 2 & Rleg < 2
tab Larm_Rleg, mi

gen cwp_m = .
replace cwp_m = 1 if Rarm_Lleg == 1 & Rarm_Lleg != . & axial_skeleton >= 1 & axial_skeleton !=
.
replace cwp_m = 1 if Larm_Rleg == 1 & Larm_Rleg != . & axial_skeleton >= 1 & axial_skeleton !=
.


tab cwp_m , mi


gen cwp_m_completed = .
foreach x in 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29{
replace cwp_m_completed = 1 if mannequin_BL_`x' != .
}
```

Bibliography

Flynn TN, Huynh E, Peters TJ, Al-Janabi H, Clemens S, Moody A, Coast J. Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. Health Economics 2015;24:258-269.

Raghunathan T.E., Solenberger P.W., Van Hoewyk J. (2002) IVEware: Imputation and Variance Estimation Software User Guide, Ann Arbor, MI, University of Michigan; http://www.isr.umich.edu/src/smp/ive/ [19 May 2008].

Van Buuren S. (2007) Multiple imputation of discrete and continuous data by fully conditional specification. Statistical Methods in Medical Research, 16, 219–242, DOI: 10.1177/0962280206074463

White IR, Carpenter J, Horton NJ. A mean score method for sensitivity analysis to departures from the missing at random assumption in randomised trials. Stat Sin. 2018 Oct;28(4):1985-2003. doi: 10.5705/ss.202016.0308. PMID: 30283213; PMCID: PMC6166859.

White, I.R. (2017). RCTMISS: Stata module to analyse a randomised controlled trial (RCT) allowing for informatively missing outcome data. Statistical Software Components. S458304, Boston College Department of Economics, revised 26 Dec 2018.