Sanderson, E., Peters, TJ., & Gooberman-Hill, R. (2017, Jun). STAR: Support and Treatment After joint Replacement Statistical Analysis Plan. Unpublished.

Evaluation of a care pathway for patients with long-term pain after knee replacement

Statistical Analysis Plan

Version 2.2 (11th June 2019)

The following people have reviewed the Statistical Analysis Plan and agree with the contents				
Name	**Role**	**Title**	**Signature**	**Date**
Emily Sanderson	**Author**	**Research Associate in Medical Statistics**	E Sanderson	10/06/19
Professor Tim Peters	**Statistical Reviewer**	**Professor of Primary Care Health Services Research**	T.J. Peters	10.6.19
Professor Rachael Gooberman-Hill	**Chief Investigator**	**Professor of Health and Anthropology**	R. Gooberman	10th June 2019

Funding

This study is funded by the National Institute for Health Research (NIHR) [NIHR Programme Grant for Applied Research programme (Grant Reference Number RP-PG-0613-20001]. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

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 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 51** 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 protocol refer to "version 9, 04-02-2019".

2. TRIAL BACKGROUND AND OBJECTIVES

2.1. Background

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

2.2. Trial objectives and aims

Please refer to the trial protocol, section titled “Aims and Objectives”.

3. TRIAL DESIGN AND PROCEDURES

3.1. Trial design and configuration

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

3.2. Trial centres

Please refer to the trial protocol, section titles “Patient selection and recruitment”.

3.3. Selection criteria

Please refer to the trial protocol, section titled “Selection criteria”.

3.4. Description of interventions

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

3.5. Control: Care as usual

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

3.6. Randomisation procedures

After patients have provided written, informed consent to participate and have completed and returned a baseline questionnaire, they will be randomly allocated to the STAR pathway or usual care. Randomisation will occur as soon as possible after the baseline questionnaire is received by the research team. Randomisation with allocation concealment will be conducted remotely via the Bristol Randomised Trials Collaboration using a web-based randomisation system. Randomisation will take place on a 2:1 basis to ensure that the intervention service is running at sufficient capacity to enable a pragmatic assessment of its effectiveness and, particularly, cost-effectiveness. If the intervention is operating to a sufficient degree of capacity per-protocol and CACE analyses will be more reliable and have higher power. To ensure reasonable balance between the two treatment groups, allocation will be minimised by pain in the replaced knee (assessed with both the Brief Pain Inventory

Severity and the Brief Pain Inventory Interference Scales – these scores are both categorised using tertiles of STAR PACE data for these scores) and stratified by orthopaedic centre. Randomisation will be performed by a member of the research team at the co-ordinating centre and the local researcher at each site will then inform participants of the result.

3.7. Sample size and justification

We estimate that 20% of patients who have had primary total knee replacement will have long-term post-surgical pain. Based on our recent trial in total knee replacement (Wylde V, 2015), we estimate conservatively that we will achieve a recruitment rate of 40% (an 8% conversion rate for number randomised out of those screened). Surgical audit data show that the four trial centres conduct over 1,900 primary total knee replacement procedures annually. Over 30 months, this equates to 4,750 procedures, and we estimate that 950 of these patients will have long-term post-surgical pain. With a recruitment rate of 40% we can recruit 380 patients over this period. In our recent trial we achieved 83% follow-up (Wylde et al, 2015); therefore, allowing for a generous 25% loss to follow-up in the STAR trial, a total of 380 participants randomised would yield 285 for analysis. For a 2:1 intervention:control randomisation ratio we would have a power of 80% to 90% to detect standardised differences of between 0.35 to 0.40 standard deviations (SDs) using a 2-sided 5% significance level. From previous studies (Dworkin RH, 2008),(Bruce J, 2014), the SDs for each of the BPI Interference and Pain Severity scales for patients with long-term post-surgical pain has been observed to be approximately 2, in which case the target effect size translates to a difference between intervention and control groups of between 0.7 and 0.8 scale points for both scales. Such a difference would be worthwhile detecting clinically, since the current consensus statement indicates that differences of approximately one scale point can be deemed the minimally important difference for both of these scales (Bruce J, 2014),(Kroenke K, 2009).

3.8. Blinding and breaking of blind

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

3.9. Trial committees

Please refer to the trial protocol, section titled “Trial organisation and oversight”.

3.10. Outcome measures

Please refer to the trial protocol, section titled “Outcome measures”.

4. GENERAL ANALYSIS CONSIDERATIONS

4.1. Analysis populations

The primary analysis of the data will be on a complete case basis, in accordance as far as possible with the intention to treat (ITT) principle whereby we analyse as randomised, disregarding protocol deviations or non-compliance. Sensitivity analyses will utilise imputation methods to handle missing data to ascertain whether their exclusion in the primary analysis has had any effect within an ITT paradigm. In addition, a crude per protocol analysis will be performed using those patients who adhered with the intervention sufficiently. Adherence to the intervention, for a patient in the intervention arm, is defined as the patient attending their assessment clinic appointment and for patients in the control arm, adherence cannot be measured as they adhere by using usual care. We assume that all patients in the control arm have adhered to the intervention allocation. This per protocol analysis will address the same points as the comparison based on groups as allocated (hereafter referred to as 'the ITT analysis') just using a different population. Since these results are likely to be biased, we will also use the Complier Average Causal Effect (CACE) approach to adjust for any selection effects in terms of adherence.

4.2. 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 which 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 questions: 1, 4, 5, 6, 8, 9 and 10. This is then standardised to range from 0 to 100 by multiplying by 3.57.

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

- **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**

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

- **EQ-5D-5L**

The EQ-5D-5L provides a state of 5 characters 'XXXXX'. Each of the five items in Section E provide an element of the state from 1 to 5. The best-case scenario is 11111 which would mean there is no problem with each area. The worst-case scenario is 55555 and this indicates that there are high levels of concern with all five areas (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). The EQ-5D-5L will be used by the Health economics team and will not be used for Statistical analysis.

- **Short Form-12**

This outcome is derived by software using responses from Section F of the questionnaire. The statistician will process the data in the software to give a score which will be in the desired format to be analysed.

- **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 H 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 I of the questionnaire and provides a state of 5 characters '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. 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 sthe following items of Section J 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 primary analysis however we will also explore analysis of the individual sub scales.

- **Pain Solutions Questionnaire**

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 K 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 the first four questions of Section L. 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 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 at 3 months is not recorded in the trial, we will classify patients based on the other elements of the definition.

- **Free text comment box**

The free text comments at the end of each questionnaire will be analysed thematically by the qualitative team and will not be used for statistical analysis other than to inform responses to other sections of the questionnaire.

4.3 Procedures for missing data

In all tables missing data will be indicated by footnotes. If the amount of missing data differs substantially between treatment groups potential reasons will be explored. Sensitivity

analyses will be conducted (including 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 ITT analyses of complete cases with ITT analyses where missing data were imputed will be presented in **Table 17 - Table 29**.

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.

In the event of missing data, we will follow guidelines where applicable and use mice to impute scores if the missingness exceeds the guidelines.

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 (ie. 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. For this trial we will fill the missing value with the mean of the completed items if one item is missing. If two or more items are missing, we will impute the whole score using mice.

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. If more than one item is missing from each subscale, we will impute the whole score using mice.

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 impute the whole score using mice.

5. DESCRIPTION OF PARTICIPANT CHARACTERISTICS

5.1. 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.

5.2 Baseline characteristics

Baseline questionnaires will be completed by patients. These data will be summarised in **Table 1 - Table 3**.

As well as the baseline outcomes, demographic variables will be summarised at baseline. These demographic variables include age, sex, marital status, living arrangement, ethnic group, a measure of deprivation and level of education. The demographic variables will be summarised by treatment group, trial centre and overall to inform us of the demographic of the trial population and to check balance between treatment groups and trial centres.

Age and sex will be summarised between patients who we screened vs randomised patients to see if the population of randomised patients reflects a similar population to those who we screened.

6. ASSESSMENT OF TRIAL QUALITY

6.1. Eligibility checks

The number of patients who were assessed and were not eligible for participation in the trial will be described in **Table 4**. The STAR trial eligibility criteria have been designed to minimise patient risk. The reason for exclusion will also be recorded in **Table 4**.

6.2. 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. This is intended to minimise bias.

6.3. Trial completion

Completion rates of questionnaires will be recorded in **Table 5** and **Table 8**. Withdrawals are summarised in **Table 9** and **Table 10** at different time points (6 months and 12 months) and the reasons for withdrawal will also be recorded in **Table 9** and **Table 10**. Patients who withdraw throughout the trial will contribute data up to the point of withdrawal providing permission is given to use previously collected data. Data missing from those patients who withdraw will be imputed in the mice analysis.

6.4. Compliance

A complier to treatment is defined as a patient in the intervention arm who attends the assessment clinic or a patient in the control arm who complies with usual care (all patients in the control arm). Since non-compliance numbers are likely to be low, we set a rule to decide if we will use CACE analysis. If compliance is close to 100% then ITT analysis will be very similar to CACE analysis and so the latter will not be (additionally) informative.

If compliance is greater or equal to 95%, we will rule out the need for using CACE analysis. For compliance between 85% and 95% we will consider carrying out CACE analysis depending on the extent and pattern of adherence. For compliance below 85% we will use CACE analysis. Compliance rate will be recorded in **Figure 2.**

Compliance is recorded in **Table 46** by site and in total.

Descriptions of the number of follow up calls that the intervention group patients receive will be presented by site and in total. The time to first follow-up call will also be summarised by site and in total. These summaries can be found in **Table 47.**

6.5. Protocol deviations

When identified, protocol deviations will be detailed on the appropriate standardised form and stored in the CRF [protocol deviation form]. The form will also be scanned and sent to the co-ordinating centre for review by the Chief Investigator. All protocol deviations will be reported to the Trial Steering Committee at meetings.
Number of protocol deviations and their nature will be recorded overall, over trial centre and over treatment group. This will be presented in **Table 11 - Table 12.**

6.6. Specify & justify changes made to the planned statistical analyses

Any adjustment to the statistical analysis plan will be logged in **Table 51.**

7. ANALYSIS OF EFFECTIVENESS

7.1 Statistical analysis

STATA version 15.1 will be used for all statistical analysis.

Continuous variables will be summarised using the mean and standard deviation (SD) (or median and inter quartile range (IQR) if the distribution is highly skewed), and categorical data will be summarised as numbers and percentages.

7.2. Summary statistics

The primary and main secondary outcomes will be summarised by mean (SD) or median (IQR) at each time point of baseline, 6 months and 12 months by intervention allocation group. The percentage of missing outcomes at each time point will be reported.

The mean BPI subscales at baseline, 6 months and 12 months will be plotted in a line graph along with confidence intervals marked at each time point. This will be split by each intervention allocation group.

A summary of the primary and secondary outcomes can be seen in **Table 13.**

7.3. Primary analysis

The primary and main secondary analyses will be conducted using the ITT principle (that is, comparisons of groups as allocated) using the appropriate regression model. Assumptions for each regression will be checked to make sure the correct method of analysis is being used.

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 as a fixed effect and baseline BPI pain scores. Estimates will be calculated of the effect that intervention has on each of the BPI scores compared with usual care. Primary analysis will be adjusted for variables which are imbalanced at baseline to see if this makes a marked difference to the unadjusted primary analysis.

Results from the primary analysis will be presented in **Table 14**.

7.4. Secondary analyses

7.4.1 Adjusting for "time to assessment clinic"

The primary analysis will be repeated including a variable which adjusts for the time between randomisation and assessment clinic. This variable will be included as a covariate in the model. The results for this can be found in **Table 15.**

7.4.2 Secondary outcomes

The secondary outcomes will be analysed using appropriate regression models in a similar manner to the primary analysis. The models will adjust for trial centre, baseline BPI pain score and also the baseline scores of the outcome for which it is modelling. A summary of the primary and secondary outcomes can be seen in **Table 13.** Further adjustment for any observed baseline imbalance will only be made for secondary outcomes in the unlikely event that such adjustment had a material effect on the results for the primary outcomes, or if there is a particular cause for concern as a consequence of a known strong relationship between a specific outcome and an individual variable demonstrating (chance) imbalance at baseline.

Results from the secondary analysis will be presented in **Table 16**.

7.5. Model assumptions and Model Fit

For the above analyses, the following assumptions must be satisfied for the regression models to provide trustworthy results.

Assumptions for **Linear regression**:

1. Continuous outcome variable.
2. One or more explanatory variables can be categorical or continuous.
3. Linear relationship between each explanatory variable vs. Pain score – to be checked using a scatter plot of the outcome variable and each explanatory variable in turn.
4. Homoscedasticity - to be checked using plots of residuals.
5. Normally distributed outcome variable – to be checked with a histogram and a Q-Q plot. If data are not normally distributed, we will consider the appropriateness of using a suitable transformation such as a log transform, bearing in mind the robustness of the standard regression model.
6. No or little multicollinearity – we will verify the associations between explanatory variables using contingency tables and correlation coefficients as appropriate.

We will also check for multi-collinearity in our logistic regression models in the same manner.

Assumptions for **Logistic regression**:

1. Dichotomous outcome variable (eg, CWP(M) status: Y/N).
2. One or more explanatory variable.
3. Independent observations (the same patient cannot be recruited twice).
4. Linear relationship between the continuous explanatory variables and the logistic transformation of the outcome variables – to be checked using a scatter plot of the outcome variable and each explanatory variable in turn.

Model fit for linear regression models will be assessed by comparing the observed values to the fitted values produced by the model. These will be presented on a plot of observed values against fitted values. If the model does not fit well, log transformations will be tested to see if this makes a difference to the quality of model fit.

7.6. Sensitivity analysis

We will investigate the influence of missing data using sensitivity analyses that make different assumptions, such as “best” and “worst” case scenarios, as well as using multiple imputation by chained equation (mice) to impute missing data.

7.6.1 Overlap

A small number of patients (≤15) who are involved in the STAR trial may also be involved in similar interventional trials and thus may influence one or both trials results. We will record which patients are involved in both trials and a sensitivity analysis will be performed as a repeat of the primary analysis on patients who are only involved in STAR.

Patient burden is an issue which will be addressed in recruitment of the STAR trial. To avoid increased withdrawal among patients involved in both trials the requirements of the patient, in terms of questionnaires and follow-up for the STAR trial, will be made very clear. Patients will also be informed that the two trials are separate trials and declining participation in one will not affect participation in the other.

Sensitivity analysis results for the primary outcomes to deal with any potential overlap will be presented in **Table 32** and **Table 33**.

7.6.2 Per protocol and CACE analysis

We propose to carry out per protocol analyses. It will only compare individuals who remained in their allocated treatment group throughout the trial. Since this analysis is likely

to be biased, we will also use the Complier Average Causal Effect (CACE) approach if the quantity of compliant patients (those patients in the intervention group who attend their initial assessment clinic and those in the control arm who continue with usual care) satisfies the rule which we state in section *6.4. Compliance.* This provides an unbiased estimate for the treatment effect for those who have complied with the treatment group allocation. Compliers would be defined as a patient who attends the assessment clinic (intervention). This approach would be justified if the characteristics of those who adhered to the comparator treatment differed from those that adhered to the intervention. Results from these analyses will be presented as in **Table 32 - Table 44**. If there is differential adherence in the two arms we will also investigate structural mean approaches as described by Fischer et al (Fischer, Goetghebeur, Vrijens, & White, 2011) and, separately, use extensions of CACE as described by White et al (White, Kalaitzaki, & Thompson, 2011).

7.6.3 Analysis accounting for time delay in assessment clinic

Since some of the patients had their assessment clinic more than 3 months after their TKR surgery we will repeat the primary analysis but restrict it to those patients who had their assessment clinic within three months of their TKR surgery as stated in the protocol.

7.6.4 Repeated measures analysis

A repeated measures analysis will be performed using a multi-level mixed effects linear regression to model each of the BPI subscales (measured at 6 months and 12 months) on the baseline measure of the outcome, stratification variables and any variables which show imbalance at baseline. The repeated measures model will also include an interaction between the intervention allocation and the time point. The p-value of the interaction term will be the focus of analysis. This analysis will account for clustering within patient using a random effect. The analysis will include those patients who have at least one non-missing outcome over the two time-points. The results of this analysis will be presented in **Table 45.**

7.7. Exploratory/other analysis

We recognise that there will be low power for subgroup and exploratory analyses and therefore only cautious conclusions will be drawn from them.

7.7.1 Exploratory analysis

Exploratory analyses such as CACE methodology will be used to estimate the effect in those patients able to comply with their allocated intervention.

7.7.2 Subgroup analysis

Subgroup analyses on the primary outcomes will be performed by introducing appropriate interaction terms between the intervention group and other patient characteristics in the regression models, to investigate any differential effects in certain subgroups of the population. These factors will be trial centre and baseline Oxford Knee Score. The OKS will be treated as a continuous variable ranging 0-48; however, for descriptive purposes we will consider the conventional categories of: 0-19 (severe knee arthritis); 20-19 (moderate to severe knee arthritis); 30-39 (moderate knee arthritis) and 40-48 (satisfactory joint function).

Another subgroup analysis will be performed for those sites who limit patients who are less than 1 hour drive away versus those who accept all patients regardless of the distance from site.

The Pain Solution (PaSol) Questionnaire measures two opposing modes of coping with pain. First, ‘assimilative coping’ captures the style in which patients increase effort and tenaciously pursue their goal of pain control. Patients scoring high on the ‘solving pain’ subscale and low on the ‘meaningful life despite pain’ and ‘acceptance of the insolubility of pain’ sub-scales are high in assimilative pain coping. Second, ‘accommodative coping’ captures the style in which patients accept that there is unlikely to be a cure for pain and a switch in goals to living with pain is a sensible next step. Patients scoring low on the ‘solving pain’ subscale and high on the ‘meaningful life despite pain’ and ‘acceptance of the insolubility of pain’ sub-scales are high in assimilative pain coping. There is a strong possibility that those with different coping profiles will respond differently to intervention. Subgroup analyses will be performed using baseline PASOL data to investigate differential effects on outcomes through the use of the relevant interaction terms in regression models. First a (continuous) composite score on a single dimension of accommodative to assimilative

scale will be created using an established algorithm (Geert Crombez, 2008) based on simple summation of scores accounting for their logical direction(s). Second, a three-category scale will be produced as follows: a) high in assimilative and low in accommodative; b) no clear preference; c) low in assimilative and high in accommodative. The exact determination of the thresholds inherent in the categorical version will be influenced both by the category labels of the underlying items/scores and their frequencies – for example, both taking into consideration any clear modality in the frequency distribution and so that none of the above three categories are either too large or too small to be of value in the analysis. It is emphasised, though, that the categorisation will be determined in advance of the regression models considering the relevant interactions, and that the primary subgroup analyses for this measure will be for the continuous version of the underlying variable.

7.7.3 Trial Centre Effect

Trial centre will be included in the regression models as a fixed effect to analyse the outcome measures. This will inform us of the effect that the trial centre has on each outcome.

8. ANALYSIS OF SAFETY

8.1. Adverse reactions

Data on adverse reactions and serious adverse reactions will also be collected and closely monitored to ensure the ongoing safety of participants. Adverse reactions will be recorded on a standardised adverse reactions report form. All serious adverse reactions will be notified to the trial sponsor (North Bristol NHS Trust) and reviewed by the Trial Steering Committee. Data on adverse reactions will be collected from trial questionnaires and during telephone contact with participants. Numbers of adverse reactions and their severity will be recorded in **Table 48**. Details of the adverse reactions will be presented in **Table 49** and **Table 50.**

Please refer to the trial protocol, section titled “Safety Reporting”

9. FINAL REPORT TABLES AND FIGURES

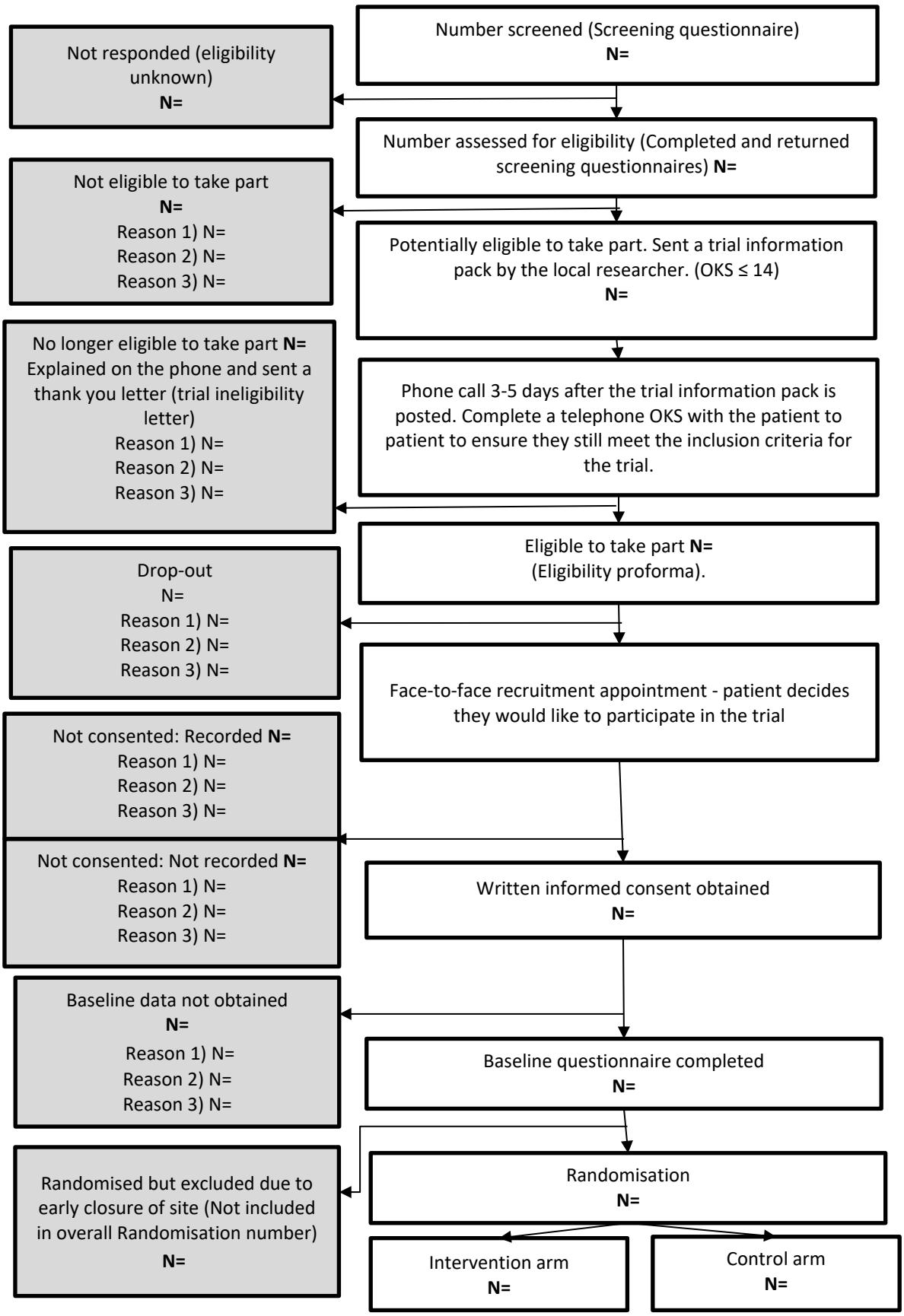


Figure 1: Consort flow diagram to monitor the number or patients included in the trial up to randomisation

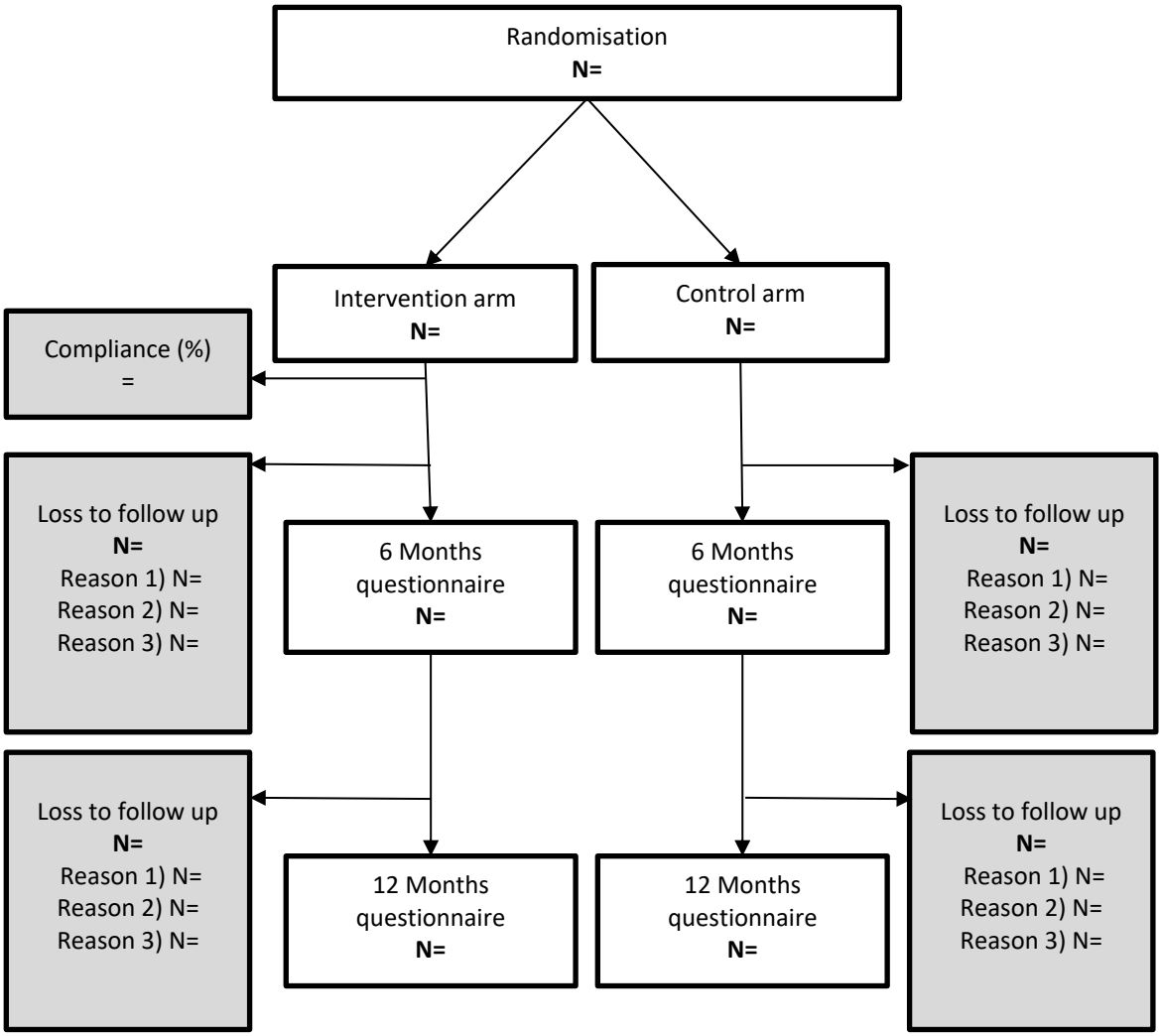
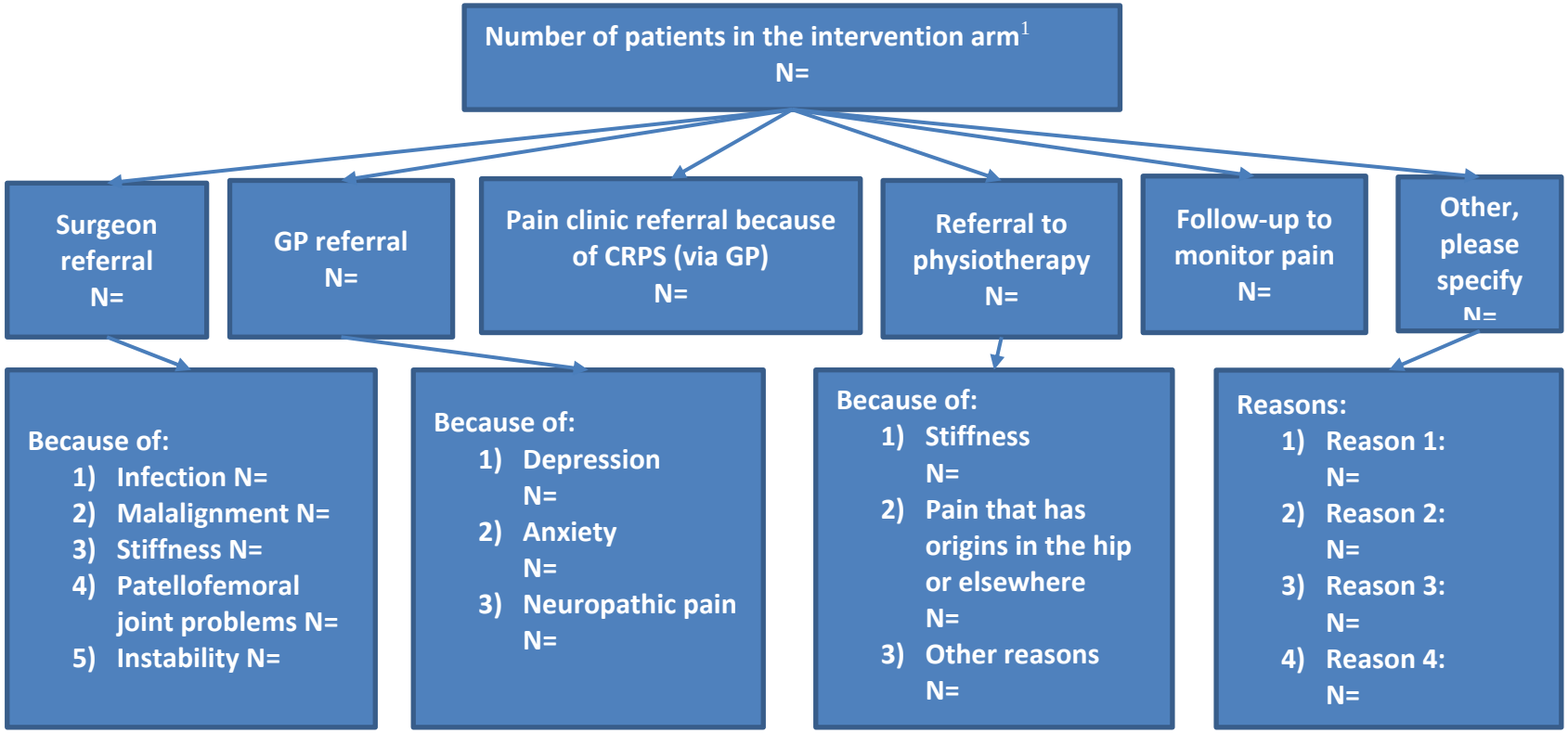


Figure 2: Consort diagram to monitor the number of patients included in the trial post randomisation

Figure 3: flow diagram to show the number of patients to receive each treatment which the intervention leads to



Number of patient receiving one referral	
Number of patient receiving two referral	
Number of patient receiving three referral	
…	

Tables

9.1. Subject characteristics and baseline summaries

Table 1: Baseline statistics for participants overall.

Demographic Variables							
	N	**# Missing**	**mean**	**s.d.**	**min**	**med**	**max**
Age							
	N	**Number of Male (%)**			**Number of Female (%)**		
Sex							
	N	**Single (%)**	**Married/ partner (%)**	**Divorced/ separated (%)**	**Widowed (%)**	**Other (%)**	
Marital Status							
	N	**Alone (%)**	**With husband/ wife/ Partner (%)**	**With Somebody else (%)**	**Other (%)**		
Living arrangement							
	N	**White (%)**	**Mixed (%)**	**Asian (%)**	**Black (%)**	**Chinese (%)**	**Other (%)**
Ethnic Group							
Outcome Measures							
	N	**# Missing**	**mean**	**s.d.**	**min**	**med**	**max**
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							
	N	**Number of "Normal" (%)**		**# "Borderline" (%)**		**# "Clinical" (%)**	
HADS: Anxiety							
HADS: Depression							
	N	**# "Rarely" (%)**	**# "Sometimes" (%)**	**# "Often" (%)**	**# "Most of the time" (%)**	**# "All of the time" (%)**	
Section A: Question 5							
Section D: Question 8							
	N	**# "Much Better" (%)**	**# "A bit better" (%)**	**# "The same" (%)**	**# "A bit worse" (%)**	**# "Much worse" (%)**	
Section L: Question 5							
	N	**# CWP(M) positive (%)**			**# CWP(M) negative (%)**		
Body Map (CWP(M))							

Table 2a: Baseline statistics for Site a (Table 2b, 2c, 2d for different sites b, c, d etc)

Demographic Variables							
	N	**# Missing**	**mean**	**s.d.**	**min**	**med**	**max**
Age							
	N	**Number of Male (%)**			**Number of Female (%)**		
Sex							

	N	Single (%)	Married/ partner (%)	Divorced/ separated (%)	Widowed (%)	Other (%)
Marital Status						
	N	Alone (%)	With husband/ wife/ Partner (%)	With Somebody else (%)	Other (%)	
Living arrangement						
	N	White (%)	Mixed (%)	Asian (%)	Black (%)	Chinese (%)	Other (%)
Ethnic Group							
Outcome Measures							
	N	# Missing	mean	s.d.	min	med	max
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							
	N	Number of "Normal" (%)		# "Borderline" (%)		# "Clinical" (%)	
HADS: Anxiety							
HADS: Depression							
	N	# "Rarely" (%)	# "Sometimes" (%)	# "Often" (%)	# "Most of the time" (%)	# "All of the time" (%)	
Section A: Question 5							
Section D: Question 8							
	N	# "Much Better" (%)	# "A bit better" (%)	# "The same" (%)	# "A bit worse" (%)	# "Much worse" (%)	
Section L: Question 5							
	N	# CWP(M) positive (%)			# CWP(M) negative (%)		
Body Map (CWP(M))							

Table 3a: Baseline statistics for patients in the intervention group (Table 3b will be be the same but for the control group)

Demographic Variables							
	N	# Missing	mean	s.d.	min	med	max
Age							
	N	Number of Male (%)			Number of Female (%)		
Sex							
	N	Single (%)	Married/ partner (%)	Divorced/ separated (%)	Widowed (%)	Other (%)	
Marital Status							
	N	Alone (%)	With husband/ wife/ Partner (%)	With Somebody else (%)	Other (%)		
Living arrangement							
	N	White (%)	Mixed (%)	Asian (%)	Black (%)	Chinese (%)	Other (%)
Ethnic Group							
Outcome Measures							
	N	# Missing	mean	s.d.	min	med	max

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							
	N	**Number of "Normal" (%)**		**# "Borderline" (%)**		**# "Clinical" (%)**	
HADS: Anxiety							
HADS: Depression							
	N	**# "Rarely" (%)**	**# "Sometimes" (%)**	**# "Often" (%)**	**# "Most of the time" (%)**	**# "All of the time" (%)**	
Section A: Question 5							
Section D: Question 8							
	N	**# "Much Better" (%)**	**# "A bit better" (%)**	**# "The same" (%)**	**# "A bit worse" (%)**	**# "Much worse" (%)**	
Section L: Question 5							
	N	**# CWP(M) positive (%)**			**# CWP(M) negative (%)**		
Body Map (CWP(M))							

9.2. Trial quality summaries

Table 4: Eligibility summary

	# screened patients	**# eligible to participate**	**Eligibility rate**	**Reasons for ineligibility**
Site 1				
Site 2				
Site 3				
Site 4				
…				
…				
Overall				

Table 5: Questionnaire completion summary over trial centres (BASELINE).

		# questionnaires completed sufficiently to produce outcome measure							
		Site 1		**Site 2**		**…**		**…**	
	# questionnaires administered	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
BPI Severity									
BPI Interference									
OKS									
PainDETECT									

DN-4									
Patient Satisfaction									
Short form-12									
HADS									
ICECAP-A									
Pain Catastrophizing scale									
Pain Solution Questionnaire									
Body Map (CWP(M))									

Table 6: Questionnaire completion summary over trial centres (6 MONTH).

		# questionnaires completed sufficiently to produce outcome measure							
		Site 1		**Site 2**		**…**		**…**	
	# questionnaires administered	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
BPI Severity									
BPI Interference									
OKS									
PainDETECT									
DN-4									
Patient Satisfaction									
Short form-12									
HADS									
ICECAP-A									
Pain Catastrophizing scale									
Pain Solution Questionnaire									
Body Map (CWP(M))									

Table 7: Questionnaire completion summary over trial centres (12 MONTH).

		# questionnaires completed sufficiently to produce outcome measure							
		Site 1		**Site 2**		**…**		**…**	
	# questionnaires administered	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
BPI Severity									
BPI Interference									
OKS									
PainDETECT									
DN-4									
Patient Satisfaction									

Short form-12									
HADS									
ICECAP-A									
Pain Catastrophizing scale									
Pain Solution Questionnaire									
Body Map (CWP(M))									

Table 8: Questionnaire completion summary treatment groups.

		# questionnaires completed sufficiently to produce outcome measure			
		Intervention		**Control**	
	# questionnaires administered	**N**	**%**	**N**	**%**
BPI Severity					
BPI Interference					
OKS					
PainDETECT					
DN-4					
Patient Satisfaction					
Short form-12					
HADS					
ICECAP-A					
Pain Catastrophizing scale					
Pain Solution Questionnaire					
Body Map (CWP(M))					

Table 9: Withdrawal summary over trial centre

	# patients randomised	**# withdrawals at 6 months N(%)**	**# withdrawals at 12 month N(%)**	**Reasons for withdrawals**
Site 1				
Site 2				
…				
…				
Overall				

Table 10: Withdrawal summary over treatment group

	# patients randomised	**# withdrawals prior to randomisation**	**# withdrawals by 6 months**	**# withdrawals by 12 month**	**Reasons for withdrawals**
Intervention					
Control					
Overall					

Table 11: Number of Protocol deviations

	Site				
	Site 1	Site 2	…	…	Total
Intervention					
Control					
Total					

Table 12: Protocol deviations

Protocol deviation	Site	Intervention/ Control

9.3. Outcome summaries

Table 13: Primary and Secondary endpoint summary

Outcome measure	Type of data	Range of values	Regression model	Efficacy parameters
BPI – Pain severity scale	Continuous	0-10 (best to worst)	Linear regression	Mean/Median/Log mean score
BPI – Pain Interference scale	Continuous	0-10 (best to worst)	Linear regression	Mean/Median/Log mean score
Oxford Knee Score (OKS)	Continuous	0-48 (worst to best)	Linear regression	Mean/Median/Log mean score
Douleur Neuropathique 4 (DN-4)	Continuous	0-7 (best to worst)	Linear regression	Mean/Median/Log mean score
PainDETECT	Continuous	-1-38 (best to worst)	Linear regression	Mean/Median/Log mean score
Hospital Anxiety and Depression Scale (HADS)	Ordinal	Each subscale: 0-21 (best to worst) normal score (0-7); borderline anxiety/depression (8-10) and clinical anxiety/depression (≥11)	Linear regression with dummy variables	Mean/Median/Log mean score
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)	Linear regression	Mean/Median/Log mean score
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)	Linear regression	Mean/Median/Log mean score
Self-Administered Patient Satisfaction Scale for Primary Hip and Knee Arthroplasty	Continuous	25-100 (worst to best)	Linear regression	Mean/Median/Log mean score
ICECAP-A	Continuous	-0.001 to 1 (worst to best)	Linear regression	Mean/Median/Log mean score
Short Form-12	Continuous		Linear regression	Mean/Median/Log mean score
Body Map	Binary	0/1: CWP(M) or not	Logistic regression	Odds ratio
Q5 Section A	Ordinal	“Rarely”, “Sometimes”, “Often”, “Most of the time”, “All of the time”	Linear regression with dummy variables	Mean/Median/Log mean score
Q8 Section D	Ordinal	“Rarely”, “Sometimes”, “Often”, “Most of the time”, “All of the time”	Linear regression with dummy variables	Mean/Median/Log mean score
Q5 Section L	Ordinal	“Much better”, “A bit better”, “The same”, “A bit worse”, “Much worse”	Linear regression with dummy variables	Mean/Median/Log mean score
Resourse use	Used by Health Economics			
EQ-5D-5L	Used by Health Economics			

9.4. Primary outcome results

Table 14: Primary outcome table

	Usual care			Intervention					
	N	Mean	SD	N	Mean	SD	Difference in means[1]	95% CI	P-value
BPI Severity									
BPI Interference									

[1] Adjusted for trial centre and baseline OKS

9.5. Secondary outcomes results

Table 15: Secondary analysis – adjusting for 'time to assessment clinic'

	Usual care			Intervention					
	N	Mean	SD	N	Mean	SD	Difference in means[1]	95% CI	P-value
BPI Severity									
BPI Interference									

[1] Adjusted for trial centre and baseline OKS and 'time to assessment clinic'

Table 16: Secondary outcomes tables

	Usual care			Intervention			Difference in means[1]	95% CI	P-value
	N	Mean	SD	N	Mean	SD			
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 OKS

9.6. 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 17:Sensitivity analysis for missing data

	N	Difference in means[a]	95% CI	p-value
Complete case				
“Best” case scenario				
“Worst” case scenario				
mice				

[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 18: Sensitivity analysis for missing data

	N	Difference in means[a]	95% CI	p-value
Complete case				
“Best” case scenario				
“Worst” case scenario				
mice				

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

9.7. Sensitivity analysis for secondary endpoints

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 Secondary outcome of **OKS**.

Table 19: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
“Best” case scenario				
“Worst” case scenario				
mice				

[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 Secondary outcome of **DN-4**.

Table 20: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
“Best” case scenario				
“Worst” case scenario				
mice				

[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 Secondary outcome of **PainDETECT**.

Table 21: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				

"Best" case scenario				
"Worst" case scenario				
mice				

[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 Secondary outcome of **Pain Catastrophizing Scale**.

Table 22: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
"Best" case scenario				
"Worst" case scenario				
mice				

[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 Secondary outcome of **Pain Solution Questionnaire (PaSol)**.

Table 23: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
"Best" case scenario				
"Worst" case scenario				
mice				

[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 Secondary outcome of **Self-Administered Patient Satisfaction Scale for Primary Hip and Knee Arthroplasty**.

Table 24: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
"Best" case scenario				
"Worst" case scenario				
mice				

[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 Secondary outcome of **ICECAP-A**.

Table 25: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
"Best" case scenario				
"Worst" case scenario				
mice				

[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 Secondary outcome of **Short form-12**.

Table 26: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
"Best" case scenario				
"Worst" case scenario				
mice				

[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 Secondary outcome of **Hospital Anxiety Scale (HADS: Anxiety)**.

Table 27: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
"Best" case scenario				
"Worst" case scenario				
mice				

[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 Secondary outcome of **Hospital Depression Scale (HADS: depression)**.

Table 28: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
Complete case				
"Best" case scenario				
"Worst" case scenario				
mice				

[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 Secondary outcome of **Chronic Widespread Pain (Body Map)**.

Table 29: Sensitivity analysis secondary endpoint results

	N	Odds Ratio[a]	95% CI	p-value
Complete case				
"Best" case scenario				
"Worst" case scenario				
mice				

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

Sensitivity analysis – Overlap of patient sample with other interventional trials

Comparison of results of ITT analysis of all cases with ITT analysis where only patients involved in STAR are analysed for primary outcome of BPI Severity scale.

Table 30: Overlap sensitivity analysis for BPI Severity scale

	N	Difference in means[a]	95% CI	p-value
Overall ITT analysis				
Only STAR Participants				

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

Comparison of results of ITT analysis of all cases with ITT analysis where only patients involved in STAR are analysed for primary outcome of BPI Interference scale.

Table 31: Overlap sensitivity analysis for BPI Severity scale

	N	Difference in means[a]	95% CI	p-value
Overall ITT analysis				
Only STAR Participants				

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

8.8. Per protocol and CACE analysis

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE analysis for primary outcome of **BPI Severity Score**.

Table 32: Sensitivity analysis for missing data

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

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

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for primary outcome of **BPI Interference Score**.

Table 33: Sensitivity analysis for missing data

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

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

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **OKS**.

Table 34: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

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

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **DN-4**.

Table 35: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

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

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **PainDETECT**.

Table 36: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

[a] Adjusted for trial centre and for baseline OKS

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **Pain Catastrophizing Scale**.

Table 37: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

[a] Adjusted for trial centre and for baseline OKS

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **Pain Solution Questionnaire (PaSol)**.

Table 38: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

[a] Adjusted for trial centre and for baseline OKS

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **Self-Administered Patient Satisfaction Scale for Primary Hip and Knee Arthroplasty**.

Table 39: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

[a] Adjusted for trial centre and for baseline OKS

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **ICECAP-A**.

Table 40: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

[a] Adjusted for trial centre and for baseline OKS

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **Short form-12**.

Table 41: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

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

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **Hospital Anxiety Scale (HADS: Anxiety)**.

Table 42: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

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

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **Hospital Depression Scale (HADS: Depression)**.

Table 43: Sensitivity analysis secondary endpoint results

	N	Difference in means[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

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

Comparison of results of ITT analysis of complete cases with per protocol analysis and CACE for Secondary outcome of **Chronic Widespread Pain (Body Map)**.

Table 44: Sensitivity analysis secondary endpoint results

	N	Odds Ratio[a]	95% CI	p-value
ITT				
Per protocol				
CACE analysis				

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

Repeated measures analysis

Table 45: Repeated measures analysis of BPI subscales

	Intervention		Usual Care				
	N groups	Average obs per group	N groups	Average obs per group	Difference in means	95% CI	P-value
BPI Pain subscale							
BPI Interference subscale							

**adjusted for:*

Table 46: Compliance

	Number of patients randomised to intervention group	Number of patients who attend intervention appointment	Compliance (%)
Site 1			

Site 2			
…			
…			
Total			

Table 47: Summaries of follow up calls received

	Site 1	Site 2	…	Total
Number of patients in the intervention arm				
Number of patients to have received at least 1 follow-up call				
Average number of follow-up calls received per intervention patient				
Average time to first follow-up call (in weeks)				

8.9. Safety results

Table 48: Reporting Adverse reactions

Relatedness to trial intervention:		Frequency
Severity:	Not Serious	
	Serious unexpected	
	Serious expected	

Table 49: Adverse reactions

Adverse reaction	Site	Intervention/ Control

Table 50: Serious Adverse reactions

Adverse reaction	Site	Intervention/ Control

Amendments to the SAP

Table 51: Amendments to the SAP

Previous version	Previous date	New version	New date	Brief summary of changes
V2.0	01.06.2017	V2.1	02.11.2017	Comments added following Oxford closure.
V2.1	02.11.2017	V2.2	11.06.2019	Added repeated measures analysis and added analysis to account for timings of clinics and subgroup analysis based on baseline PaSol.

10. APPENDICES

10.1. Stata code for derived variables

```
*BPI_severity

gen bpi_severity = (worst_bl+least_bl+average_bl+rightnow_bl)/4


*BPI_interference

gen bpi_int =
(interfere_gen_bl+interfere_mood_bl+interfere_walk_bl+interfere_norm_bl+interfere_relation_bl+
interfere_sleep_bl+interfere_life_bl)/7


*Oxford knee score

gen oks =
replaced_pain_bl+replaced_wash_bl+replaced_car_bl+replaced_walk_bl+replaced_sat_bl+replaced_li
mp_bl+replaced_kneel_bl+replaced_trouble_bl+replaced_work_bl+replaced_giveway_bl+replaced_shop
_bl+replaced_stairs_bl
* OKS pain subscale
oks_pain_raw =
replaced_pain_bl+replaced_walk_bl+replaced_sat_bl+replaced_limp_bl+replaced_trouble_bl+replace
d_work_bl+replaced_giveway_bl+
oks_pain_sta = 3.57*oks_pain_raw
*OKS function subscale
oks_func_raw =
replaced_wash_bl+replaced_car_bl+replaced_kneel_bl+replaced_shop_bl+replaced_stairs_bl
oks_func_sta = 5*oks_func_raw


*DN-4

*gen yn_feelpain_burn_bl = .
*replace yn_feelpain_burn_bl = 1 if feelpain_burn_bl = "yes"
*replace yn_feelpain_burn_bl = 0 if feelpain_burn_bl = "no"

*gen yn_feelpain_cold_bl = .
*replace yn_feelpain_cold_bl = 1 if feelpain_cold_bl = "yes"
*replace yn_feelpain_cold_bl = 0 if feelpain_elect_bl = "no"

*gen yn_feelpain_elect_bl = .
*replace yn_feelpain_elect_bl = 1 if feelpain_elect_bl = "yes"
*replace yn_feelpain_elect_bl = 0 if feelpain_elect_bl = "no"

*gen yn_painfeel_tingling_bl = .
*replace yn_painfeel_tingling_bl = 1 if painfeel_tingling_bl = "yes"
*replace yn_painfeel_tingling_bl = 0 if painfeel_tingling_bl = "no"

*gen yn_painfeel_pins_bl = .
*replace yn_painfeel_pins_bl = 1 if painfeel_pins_bl = "yes"
*replace yn_painfeel_pins_bl = 0 if painfeel_pins_bl = "no"

*gen yn_painfeel_numbness_bl = .
*replace yn_painfeel_numbness_bl = 1 if painfeel_numbness_bl = "yes"
*replace yn_painfeel_numbness_bl = 0 if painfeel_numbness_bl = "no"

*gen yn_painfeel_itching_bl = .
*replace yn_painfeel_itching_bl = 1 if painfeel_itching_bl = "yes"
*replace yn_painfeel_itching_bl = 0 if painfeel_itching_bl = "no"

*egen dn_4 =
yn_feelpain_burn_bl+yn_feelpain_cold_bl+yn_feelpain_elect_bl+yn_painfeel_tingling_bl+yn_painfe
el_pins_bl+yn_painfeel_numbness_bl+yn_painfeel_itching_bl

egen dn_4 =
feelpain_burn_bl+feelpain_cold_bl+feelpain_elect_bl+painfeel_tingling_bl+painfeel_pins_bl+pain
feel_numbness_bl+painfeel_itching_bl
```

```
*PainDETECT

gen feelpain_replaced_bl_score = .
replace feelpain_replaced_bl_score = 0 if feelpain_replaced_bl == 1
replace feelpain_replaced_bl_score = -1 if feelpain_replaced_bl == 2
replace feelpain_replaced_bl_score = 1 if feelpain_replaced_bl == 3
replace feelpain_replaced_bl_score = 1 if feelpain_replaced_bl == 4

gen pain_detect = .
replace pain_detect =
(feelpain_sting_bl+feelpain_prick_bl+feelpain_touch_bl+feelpain_shock_bl+feelpain_temp_bl+feel
pain_numb_bl+feelpain_press_bl+feelpain_replaced_bl_score+feelpain_radiate_bl)


*HADS

gen hads_a =
(mood_wound_bl+mood_fright_bl+mood_worry_bl+mood_relax_bl+mood_butterfly_bl+mood_restless_bl+m
ood_panic_bl)

gen hads_d =
(mood_enjoy_bl+mood_laugh_bl+mood_cheerful_bl+mood_slow_bl+mood_appear_bl+mood_lookforward_bl+
mood_book_bl)

*Pain Catastrophizing Scale


gen pcs_r = (pain_away_bl+pain_mind_bl+pain_hurts_bl+pain_stop_bl)

gen pcs_m = (pain_worse_bl+pain_events_bl+pain_serious_bl)

gen pcs_h =
(pain_worry_bl+pain_can_go_on_bl+pain_terrible_bl+pain_awful_bl+pain_stand_more_bl+pain_intens
ity_bl)


*PaSol

gen pa_sol_solve =
deal_pain_search_bl+deal_pain_rid_bl+deal_pain_solut_bl+deal_pain_without_bl

gen pa_sol_meaning =
deal_pain_meaningful_bl+deal_pain_wayout_bl+deal_pain_live_bl+deal_pain_best_bl+deal_pain_way_
bl

gen pa_sol_accept = deal_pain_no_solution_bl+deal_pain_cntrl_bl+deal_pain_accept_bl

gen pa_sol_belief = deal_pain_conf_bl+deal_pain_treat_bl

*ICECAP-A



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
gen
satisfaction_scale=(satisfied_surgery_bl+satis_improve_bl+satis_housework_bl+satis_leisure_bl)
/4


*EQ-5D-5L
*Used by KG for health econ
```

*ShortForm-12
*software calculated scores

11. Bibliography

Bruce J, T. A., Powell R, Johnston M, Wells M, Heys SD, et al. (2014). Psychological, surgical, and sociodemographic predictors of pain outcomes after breast cancer surgery: A population-based cohort study. *Pain., 155*(2), 232-243.

Dworkin RH, T. D., Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. (2008). Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *The journal of pain : official journal of the American Pain Society., 9*(2), 105-120.

Fischer, K., Goetghebeur, E., Vrijens, B., & White, I. R. (2011). A structural mean model to allow for noncompliance in a randomized trial comparing 2 active treatments. *Biostatistics, 12*(2), 247-257.

Kroenke K, B. M., Damush TM, Wu J, Hoke S, Sutherland J, et al. (2009). Optimized antidepressant therapy and pain self-management in primary care patients with depression and musculoskeletal pain: a randomized controlled trial. *JAMA, 301*(20), 2099-2110.

Marques EM, B. A., Lenguerrand E, Wylde V, Noble SM. (2015). Local anaesthetic wound infiltration in addition to standard anaesthetic regimen in total hip and knee replacement: long-term cost-effectiveness analyses alongside the APEX randomised controlled trials. *BMC Med*, 13:151.

White, I. R., Kalaitzaki, E., & Thompson, S. G. (2011). Allowing for missing outcome data and incomplete uptake of randomised interventions, with application to an Internet-based alcohol trial. *Statistics in Medicine, 30*(27), 3192-3297.

Wylde V, L. E., Gooberman-Hill R, Beswick AD, Marques E, Noble S, et al. (2015). Effect of local anaesthetic infiltration on chronic postsurgical pain after total hip and knee replacement: the APEX randomised controlled trials. *Pain., 156*(6), 1161-1170.

G. Crombez, C. Eccleston, G. Van Hamme, P. Vlieger. (2008). Attempting to solve the problem of pain: a questionnaire study in acute and chronic pain patients. *Pain.*, 137 556-563