

Improving the self-management of chronic pain: COPERS

Statistical Analysis Plan

Version: 1.0
Date: 03/10/2013

Person(s) contributing to the analysis plan	
Name(s) and position(s)	Brennan Kahan Dawn Carnes Kate Homer Karla Diaz-Ordaz Miland Joshi Stephen Bremner Sandra Eldridge Stephanie Taylor Martin Underwood
Authorisation	
Position	Chief or principal investigator
Name	Professor Stephanie Taylor
Signature	
Date	
Position	Senior trial statistician
Name	Professor Sandra Eldridge
Signature	
Date	
Position	Trial Steering Committee statistician
Name	Dr. Obi Ukoumunne
Tick once reviewed	
Date	

1. INTRODUCTION

Purpose of statistical analysis plan

The purpose of this document is to provide details of the statistical analyses and presentation of results to be reported within the principal paper(s) of the COPERS trial. Subsequent papers of a more exploratory nature (including those involving baseline data only) will not be bound by this strategy but will be expected to follow the broad principles laid down within it. Any exploratory, post-hoc or unplanned analyses will be clearly identified as such in the respective study analysis report.

The structure and content of this document provides sufficient detail to meet the requirements identified by the International Conference on Harmonisation (ICH) and PCTU_SOP_SP 01_Statistical Analysis.

The ethics application was submitted in Feb 2011, and approval was granted on 18/03/2011.

The following were reviewed in preparation for a preliminary version of this document:

- ICH E9 Guidance on statistical principles for clinical trials
- ICH E3 Structure and content of clinical study reports
- CONSORT guidelines for the reporting of randomised trials

Stephen Bremner was responsible for the original statistical analysis strategy in the protocol. Brennan Kahan and Karla Diaz-Ordaz have written the statistical analysis plan under the direction of Sandra Eldridge. Dawn Carnes, Kate Homer, Martin Underwood, and Stephanie Taylor have also contributed to the writing of this statistical analysis plan. Sandy Smith has designed the database to collate and store the data from the questionnaires.

This document has been developed prior to examination of unblinded trial data. This plan is intended not to change or contradict the general aims of the protocol, but rather expand on them. In the event of a discrepancy the analyses described here will supersede those in earlier documents.

2. STUDY OBJECTIVES AND ENDPOINTS

2.1 Primary Objective

- To test the effectiveness and cost-effectiveness of a group self-management courses for people with persistent pain.

2.2 Secondary objectives

- To test the hypothesis that treatment effectiveness is moderated by baseline self-efficacy
- To test the hypothesis that long-term (12 month) effectiveness is mediated by change in self-efficacy between baseline and three months.

2.3 Outcome measures

Primary outcome

The primary outcome is the disability subsection of the Chronic Pain Grade questionnaire (CPG disability) (Von Korff, 1992) at 12 months post randomisation.

This outcome is a composite of three questions assessing the extent to which the participant's pain has interfered with or changed their ability to perform their daily activities, work, or take part in recreational, social, and family activities in the previous six months. Each of the three questions is rated on a scale of 0-10,

with 0 reflecting no change or interference, and 10 reflecting extreme change or interference.

The primary outcome is the mean of these three questions, multiplied by 10; i.e. if X1, X2, and X3 represent the three questions, and Y represents the primary outcome, then $Y=10*(X1+X2+X3)/3$. The primary outcome is therefore recorded on a scale from 0-100, with higher scores reflecting larger interference or change in the participant's ability to perform daily activities, work, or take part in recreational, social, and family activities.

Secondary outcomes

- 1) CPG disability at 6 months post randomisation
- 2) CPG pain intensity score at 6 and 12 months post randomisation.
- 3) PSEQ (Pain Self-Efficacy Questionnaire) score at 6 and 12 months post randomisation
- 4) HADS (Hospital Anxiety and Depression Scale) Anxiety score at 6 and 12 months post randomisation
- 5) HADS (Hospital Anxiety and Depression Scale) Depression score at 6 and 12 months post randomisation
- 6) CPAQ (Coping Pain and Acceptance Questionnaire) score at 6 and 12 months post randomisation
- 7) HEIQ (Health Education Impact Questionnaire) Social integration score at 6 and 12 months post randomisation
- 8) EQ-5D at 6 and 12 months post randomisation
- 9) Census global health question at 6 and 12 months post randomisation
- 10) Total Defined Daily Doses (Total DDD) consumed of psychotropic drugs up to 12 months post-randomisation
- 11) Total DDD consumed of analgesics (including all opioids and other CNS drugs) for pain up to 12 months post randomisation
- 12) Total DDD consumed of weak opioids up to 12 months post randomisation
- 13) Total DDD consumed of strong opioids up to 12 months post randomisation
- 14) Proportion of participants using weak opioids at 12 months post randomisation (defined as having received a prescription for a weak opioid up to twelve weeks before the 12 month follow-up date)
- 15) Proportion of participants using strong opioids at 12 months post randomisation (defined as having received a prescription for a strong opioid up to twelve weeks before the 12 month follow-up date)

A guide to how outcomes are derived is available in the appendix.

3. STUDY METHODS

3.1 Overall study design and plan

Target for randomisation: 391 intervention and 294 control participants

Date of first randomisation: 6th September 2011

Date of last randomisation: 18th July 2012

Trial design: Individually randomized, parallel group

Blinding: It was not possible to blind participants. Data entry and telephone follow-up are blinded

Randomised Interventions: Intervention with usual care vs Modified attention control (relaxation) with usual care

Target allocation ratio: 1:1.33 (control: intervention)

3.2 Selection of study population

Inclusion criteria:

- Adults (aged 18 or over) with chronic musculoskeletal pain

The International Association for the Study of Pain (IASP) defines chronic pain as that which has

persisted beyond normal tissue healing time - usually interpreted as three months (IASP 1986).

Examples include osteoarthritis, any chronic musculoskeletal pain, chronic widespread pain and

fibromyalgia; we excluded inflammatory arthritis such as rheumatoid arthritis. We included people with chronic pain and a past history of cancer where the chronic pain arose from non-malignant causes.

Exclusion criteria:

- Inability to give informed consent
- Not fluent in English
- Serious active co-morbidity that is more disabling to the individual than chronic pain
- Serious mental health issues that would make it difficult for an individual to participate in the group course
- People with a life expectancy of less than six months
- Substance misuse that would make it difficult for an individual to participate in the group course
- People with chronic pain arising from malignant disease because this requires specific management

3.3 Method of treatment assignment and randomisation

Participants were assigned to the intervention or control group in a 1.33 to 1 ratio (intervention: control) using stratified permuted blocks with randomly varying block lengths of 7 and 14. Site of recruitment was used as a stratification factor. Treatment assignments were carried out via a remote computerized randomisation service.

3.4 Treatment masking (blinding)

All parties were blind to allocation up to the point of randomisation and all baseline data were collected by self-completed questionnaire prior to randomisation.

After allocation, we could not blind researchers to participants' treatment allocation in their own location.

Follow-up data collected by telephone by trial research personnel was blind to treatment allocation (the London team collected data from Warwick site participants and vice versa). Because participants were aware of their treatment allocation, we used a standardised script asking participants not to divulge their allocation to the data collector. All other data were collected by self-completed

questionnaires and / or electronic databanks returned to the trial team for data entry.

The statistician analysing the data will not be blinded once any information on allocation has been received. As far as possible, data cleaning and checking by the statistician will be completed prior to information about which participants are in the control group and which in the intervention group being disclosed to them.

3.5 Sample size

The sample size calculation was based on detecting a standardised mean difference of 0.3 in pain related disability between intervention and control groups, with a power of 80% at the 5% significance level. This effect size was commensurate with the largest change seen in a recent systematic review of expert patient programmes[2], and also with the sort of change effected by interventions for other chronic pain syndromes, such as low back pain, on any continuous outcome measure[3]. A simple sample size calculation indicated that we would require data on 350 subjects. We inflated the sample size because of the possibility of a ‘clustering’ effect in the group intervention arm and chose the ratio between intervention and control participants to increase statistical efficiency [4]. Using an intra-cluster correlation coefficient (ICC) of 0.1, and assuming on average nine individuals providing data from each group results in 480 individuals needed with 275 in the intervention group and 205 in control the control group (1.33:1 intervention:control). Allowing, conservatively, for a 30% loss to follow-up (from an average of 13 individuals recruited per group) we sought to randomise 685 participants (391 intervention participants and 294 controls).

3.6 Trial Consent

Consent was gained for: participating in the trial, audio-recording the intervention sessions and accessing medical records at 12 months.

4. DATA COLLECTION

Data were collected at four time points: baseline, 12 weeks, 6 months and 12 months post randomisation. All data were collected via postal self-report questionnaires, except for data about participant co-morbidities and use of pain related medication, which were obtained from the participant's GP record.

Recruitment began in August 2011, finished in July 2012; follow-up was completed in August 2013.

4.1 Baseline data collection

Descriptive data:

- Age
- Gender (Male/Female)
- Ethnicity (White, Black or Black British, Asian or Asian British, Mixed, Other)
- English language fluency (Fluent, Good, Below Average, Poor)
- Age at which formal education ended (no formal education, age 12 or less, age 13 to 16, age 17 to 19, age 20 or over, still in full time education, other)
- Employment status (employed, unemployed and looking for work, at school or in full time education, unable to work due to long term sickness, looking after home or family, retired from paid work, other)
- Number of body systems affected by co-morbid conditions (musculoskeletal, cardiovascular, tegumental, gastrointestinal, genitourinary, mental health, ENT/optical, respiratory, neurological, endocrine/metabolic/immune, other)
- Time kept from usual activities due to pain in last 6 months (0-6 days, 7-14 days, 15-30 days, 31 or more days)
- Site recruited (London, Warwick)
- Duration of pain (0-3 months, 4-12 months, 13 months to 2 years, 3-4 years, 5-6 years, 7-10 years, more than 10 years)
- Living arrangements (lives alone, lives with others)
- Overall CPG score

Outcomes measured at baseline:

- CPG disability
- CPG pain intensity
- HADS Anxiety
- HADS Depression
- EQ-5D
- PSEQ
- CPAQ
- HEIQ
- Census global health question
- Total amount of drugs taken above the DDD in three months prior to randomization (psychotropic, weak opioids, strong opioids, analgesics)
- Opioids prescriptions (strong and weak opioids)

4.2 Twelve weeks data collection

- PSEQ

4.3 Six months data collection

- CPG disability
- CPG pain intensity
- HADS Anxiety
- HADS Depression
- EQ-5D
- PSEQ
- CPAQ
- HEIQ
- Census global health question
- Private healthcare use during previous 6 months (In addition to the core seven questionnaires (4.1 (b) above) at six months we also asked participants about their non-NHS health care resource use: private health care hospital stays, private tests, private consultations, privately purchased prescriptions/meds and devices and expenditure on social support such as transport and home help in the previous 6 months)

4.4 Twelve months data collection

- CPG disability
- CPG pain intensity
- HADS Anxiety
- HADS Depression
- EQ-5D
- PSEQ
- CPAQ
- HEIQ
- Census global health question
- Private healthcare use during previous 6 months (see 6 month data collection for details)
- Total amount of drugs taken above the DDD during follow-up (psychotropic, weak opioids, strong opioids, analgesics)
- Opioids prescriptions (strong and weak opioids)
- Other courses or activities attended during follow-up outside of COPERS trial: pain management course, expert patient programme or other self management course, other wellness or wellbeing course, return to work course, frequency of relaxation techniques (never, rarely, daily, weekly, monthly)

5. GENERAL ISSUES FOR STATISTICAL ANALYSIS

5.1 Blinding of the statistical analysis

Analysis cannot be blinded because of the allocation ratio. As far as possible all cleaning and checking of the data will be done before the statistician has access to the allocation codes.

5.2 Database

We will use a Microsoft™ Access 2007 bespoke database incorporating SQL and VBA programming code developed by PCTU.

Data quality

Single data entry was performed. 100% data entry check was performed for the primary outcome (CPG disability), EQ-5D, and randomisation code. This was performed by somebody other than the person who entered the data, and involved checking the values entered on the database matched the questionnaire. A subset (approximately 10%) of questionnaires at baseline, 12 weeks, 6 months, and 12 months were checked by comparing the values entered on the database to the questionnaire.

Database lock

Once the trial team has completed all data entry and checking, the database will be date stamped and transferred to a read-only location on the appropriate server. The statistician responsible for the analysis will conduct or oversee additional data checks. Any necessary changes will be communicated to the appropriate member of the data management team as detailed in PCTU SOP PCTU_DM_04. This process will be repeated until the statistician and data management team are

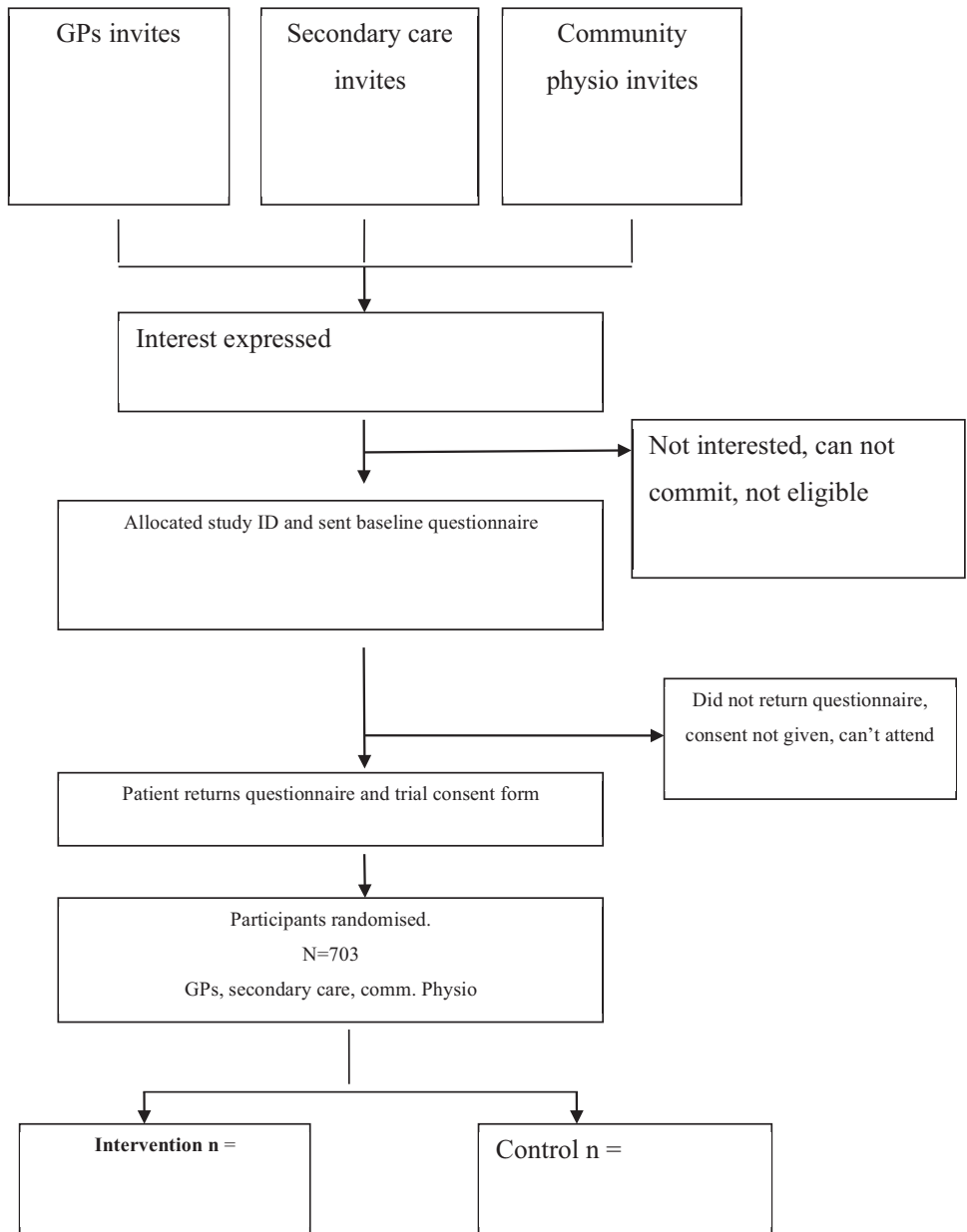
satisfied that all identifiable errors have been corrected. At this point, the database will be locked by removing access rights. After database lock, the database will be date stamped and transferred to a read-only location on the appropriate server. This dataset will be used for analysis. The database will not be locked until version 1.0 of the Statistical Analysis Plan has been finalised and signed off.

5.3 Analysis software

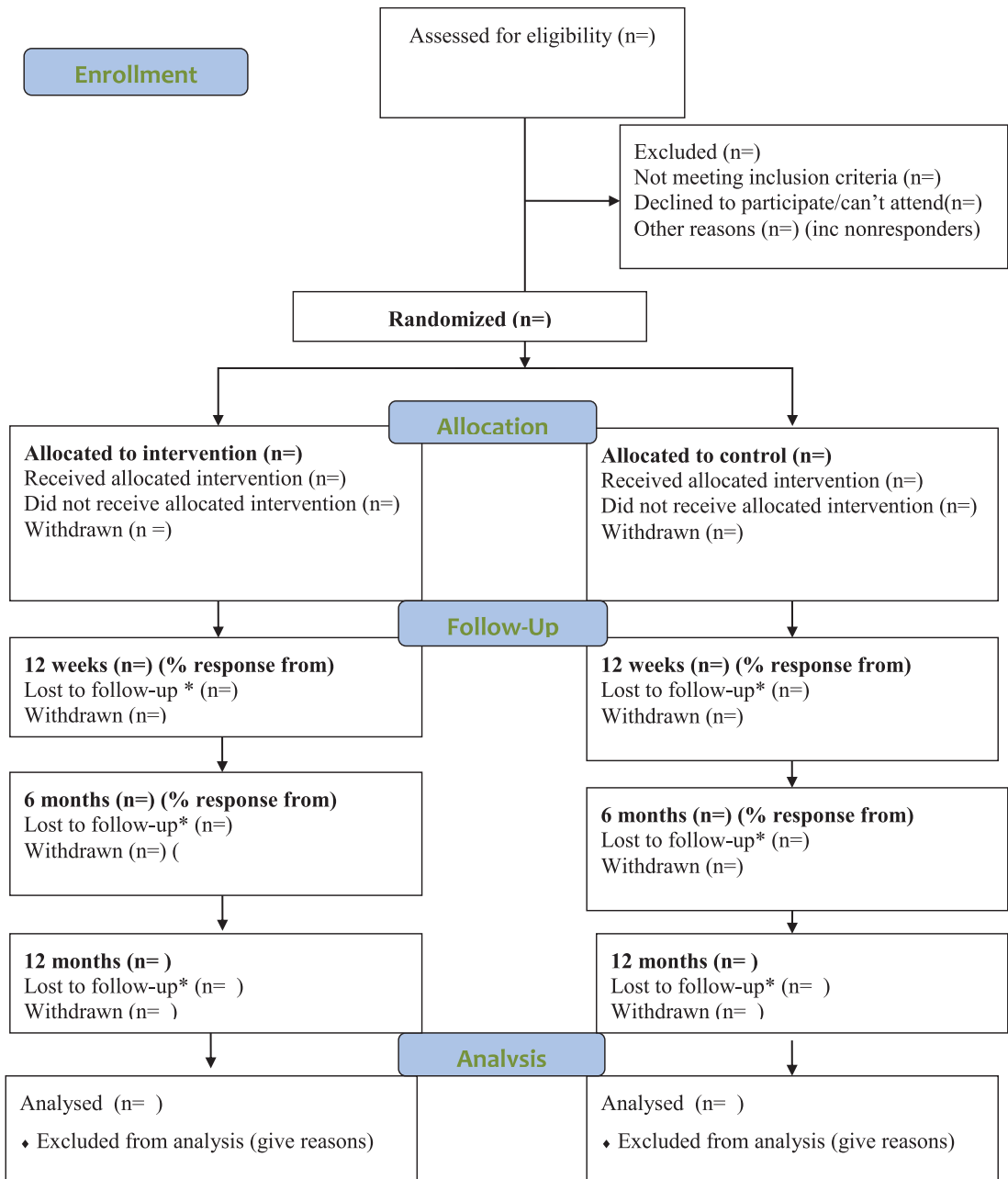
The analysis will be carried out using Stata version 12. Other packages such as R, SAS, or REALCOM may be used if necessary.

6. DESCRIPTIVE ANALYSES OF TRIAL

6. 1 Recruitment flow chart



6. 2 COPERS CONSORT Flow Diagram



*Loss to follow up = moved or phone number changed

6.3 Representativeness of sample

The age, gender and ethnic profile of randomised participants will be examined to see if they are typical of the UK population with chronic pain. For example one UK population survey showed the age in a sample of people with chronic pain to have a mean and SD of 55 and 16.7, and the proportion of males being 41% (399/966) (Parsons et al 2007). We will also examine the gender of participants who expressed an interest in the trial (were assigned a study ID), but were not randomised

Certain baseline characteristics will be compared between participants who were lost to follow-up vs. other participants (table 9). Overall numbers lost to follow-up will be included in the CONSORT flow chart.

7. ANALYSIS PRINCIPLES

7.1 GENERAL ANALYSIS PRINCIPLES

The main analysis for each outcome will use intention-to-treat (ITT) principles, meaning that all participants with a recorded outcome will be included in the analysis, and will be analysed according to the treatment group to which they were randomised. More information on which participants will be included in each analysis is available in sections 7.2 and 7.4. All p-values will be two sided, and the significance level is set at 5%.

Analyses for all outcomes will be presented as:

- The number of participants included in the analysis, by treatment group;
- A summary measure of the outcome, by treatment group (e.g. mean (SD) for continuous outcomes, number (%) for binary outcomes, etc). Only participants with a completely recorded outcome will be used to calculate the summary measure (e.g. participants who complete only 1 of 3 components of the CPG disability score will not be included in the calculation of the summary measure);
- A treatment effect, with a 95% confidence interval;
- A two-sided p-value.

All analyses will account for clustering by course in the intervention arm.

Participants in the control arm (who do not attend courses), will act as their own cluster (i.e. each participant in the control arm will belong to a 'course' where they are the only member).

Site of recruitment (London or Midlands), age, gender, and the HADS depression score at baseline will be included as covariates in each analysis. Additionally, for continuous outcomes (CPG disability, CPG pain intensity, PSEQ, HADS Anxiety, HADS Depression, CPAQ, HEIQ, and EQ-5D), the outcome measured at baseline will be included in the analysis. Continuous covariates (age, HADS depression

score, outcome measured at baseline) will be assumed to have a linear relationship with the outcome.

7.2 Primary analysis

The primary outcome (CPG disability at 12 months) will be analysed using a mixed-effects linear regression model, with 'course' as a random effect.

Restricted maximum likelihood (REML) will be used. The model will include site of recruitment, age, gender, HADS depression score, and CPG disability at baseline as covariates.

All participants who completed at least one of the three questions which form the CPG disability score at either 6 or 12 months will be included in the analysis. Participants who did not fill out any portion of the CPG disability score at either 6 or 12 months will be excluded from the analysis. It should be noted that CPG disability will be analysed separately at 6 and 12 months.

Multiple imputation (MI) will be used to account for participants who have an observed outcome at 6 months, but are missing the outcome at 12 months, as well as participants who completed some, but not all, of the questions on the CPG disability score at 12 months. 20 imputations will be performed, and results will be combined using Rubin's Rules. Only participants who will be included in the analysis will be included in the imputation model. Imputation will be performed separately within each treatment arm. The imputation model will include the three questions which form the CPG disability score at baseline, 6 months, and 12 months, as well as site of recruitment, age, gender, the HADS depression score at baseline, and employment status (employed or in full time education vs not employed or in full time education) (14 variables in total). In the intervention arm, multilevel imputation will be performed, with 'course' included in the imputation model as a random effect.

Missing data in any of the covariates to be adjusted for in the analysis (site of recruitment, age, gender, HADS depression score, CPT disability and baseline) will be accounted for using the same multiple imputation model as above.

Sensitivity analyses

Method of accounting for missing data

We will perform three sensitivity analyses for the primary outcome to assess the robustness of the results to other methods of account for missing data. The first sensitivity analysis involves specifying a different imputation model than that used in the primary analysis, and the last two sensitivity analyses involve re-analyse the primary outcome using two approaches which are not based on MI.

- We will determine which baseline covariates are associated with loss to follow-up, and include them in the imputation model. The analysis model will be the same as that described in 7.2, except for the inclusion of additional covariates in the imputation model.
- We will perform a complete case analysis, where all participants who did not complete all components of the CPG disability score at 12 months will be excluded from the analysis. The analysis model will be the same as that described in 7.2, except missing baseline covariates will be replaced using mean imputation.
- We will analyse the three components which form the CPG disability score at 12 months, rather than the CPG disability score itself. This will be done by performing a multivariate analysis, where each of the three components from the 12 month score are included in the model as outcomes (i.e. each participant will have three outcomes). A three-level mixed-effects model will be used, with random effects for ‘course’ and for participant. Treatment-by-question interactions will be included, allowing the treatment effect to vary for each of the three components. An overall treatment effect for CPG disability at 12 months will be estimated using the *lincom* function in Stata to combine the treatment estimates from the three separate components. As above, missing baseline covariates will be replaced using mean imputation.

Participants with no completed follow-ups

The primary analysis has assumed that the excluded participants (those not completing any questions on the CPG disability questionnaire at both 6 and 12 months) were missing at random (i.e. they were missing based on the covariates included in the analysis model). To assess the robustness to departures from this assumption, the primary outcome will be assessed under a range of missing-not-at-random scenarios. This will be done using the formula $\Delta = \Delta_{\text{primary}} + Y_1P_1 - Y_2P_2$, where Δ is the treatment effect under the missing-not-at-random scenario, Δ_{primary} is the treatment effect from the primary analysis, Y_1 and Y_2 are the assumed mean responses for participants with missing data in treatment groups 1 and 2 respectively, and P_1 and P_2 are the proportion of participants who were excluded from the analysis in groups 1 and 2 respectively. The standard error for Δ is assumed to be approximately equal to the standard error for Δ_{primary} . Y_2 will be varied between 10, 25, 50, 75, and 90, and for each value of Y_2 , Y_1 will be set to $Y_2 - 10$, Y_2 , and $Y_2 + 10$. For example, for $Y_2 = 25$, Y_1 will vary between 15, 25, and 35.

Re-definition of primary endpoint

The primary outcome is a composite of three questions. The first question (Q1) assesses to what extent the participant's pain has interfered with daily activities in the previous six months. This is assessed on a scale of 0-10, with higher scores indicating more interference. The last two questions assess to what extent the participant's pain has changed their ability to (a) take part in recreational, social, and family activities (Q2); and (b) work (Q3). Both these questions are measured on a scale from 0-10, with higher scores indicating more extreme change.

For the last two questions, higher change scores are meant to represent a higher *negative* change, however it is possible that some participants have misinterpreted this, and have recorded a high score to indicate a large positive change. We will therefore perform a sensitivity analysis by redefining the outcome for participants

whose scores indicate they may have misinterpreted the intended direction of the questions relating to change.

For participants with a score of 2 or less for Q1 (indicating very little interference in daily activities) *and* a score of 8 or higher on either Q2 or Q3 (intending to indicate an extreme negative change in their ability to take part in social activities or to work), we will assume the participant has misinterpreted the intended direction of the scale for Q2 or Q3 (as it is inconsistent for the pain to have had very little interference in daily activities, and for there to have been an extreme negative change in the participant's ability to take part in activities or work). We will therefore rescore Q2 or Q3 based on a reverse scale (i.e. a score of 10 will be rescored as 0, 9 will be rescored as 1, and 8 will be rescored as 2). We will then re-analyse the outcome using the same method as for the main analysis

7.3 Subgroup analyses

Subgroup analyses will be performed for the primary outcome (CPG disability at 12 months). All subgroup analyses will be performed using the same analysis model as for the primary outcome, but will also include the subgroup of interest and a treatment-by-subgroup interaction. Interaction tests will be considered significant at the 5% level. No correction will be made for multiple tests.

The following subgroups will be assessed:

(i) Non-pain:

- Co-morbidity: ≤ 3 vs. > 3 co-morbidities, including musculoskeletal
- Living arrangements: living alone vs. living with others
- Baseline self-efficacy: PSEQ score 0-20 (not likely to be confident) vs. 21-39 (more likely to be confident and to self manage) vs. ≥ 40 (confident) (Nicholas 2006, 2007)
- Socioeconomic status (SES) (based on Index of Multiple Deprivation 2010, calculated from participant postcodes via GIS: lower social class (less than observed median in data) vs higher social class (equal or greater than observed median in data))

(ii) Pain-related:

- Pain duration: 0-12 months vs 13 months to 4 years vs 5 or more years
- Baseline pain intensity: CPG intensity score 0-3 (low) vs 4-7 (medium) vs 8-10 (high)
- Baseline pain-related disability: CPG disability score 0-3 (low) vs 4-7 (medium) vs 8-10 (high)
- Baseline depression: HADS depression score <11 vs ≥ 11

7.4 Analysis of secondary outcomes

CPG disability at 6 months

This outcome will be analysed using the same methods as CPG disability at 12 months.

CPG pain intensity, HADS Anxiety, HADS Depression, and HEIQ at 6 and 12 months

These outcomes will be analysed using the same methods as CPG disability at 6 and 12 months.

PSEQ at 6 and 12 months

This outcome will be analysed using the same methods as CPG disability at 6 and 12 months, except the individual components of the PSEQ score at 12 weeks will also be included in the imputation model.

CPAQ at 6 and 12 months

This outcome will be analysed using the same methods as CPG disability at 6 and 12 months, with the exception of how CPAQ at baseline is included in the MI model. CPAQ is a composite of 20 questions – including each of these questions at each time point in the imputation model would lead to 60 variables being included (20 questions at baseline, 20 at 6 months, and 20 at 12 months) which may cause problems. We will therefore include only the individual questions for CPAQ at 6 and 12 months in the imputation model, and include the full CPAQ score at baseline (leading to 41 variables rather than 60). For participants who are missing CPAQ at baseline, we will use mean imputation.

EQ-5D at 6 and 12 months

The EQ-5D will be analysed using the same analysis model as the primary outcome (i.e. mixed-effects linear regression model, with course as a random effect, adjusted for site of recruitment, age, gender, HADS depression score, and EQ-5D at baseline).

All participants who fully complete the EQ-5D score at either 6 or 12 months will be included in the analysis. EQ-5D scores with missing components will be regarded as completely missing.

MI will be used to account for participants who are missing the outcome at either 6 or 12 months. The MI strategy will be the same as that for the primary and other secondary outcomes, except instead of imputing the individual components of the EQ-5D score, we will impute the whole score.

Census global health question at 6 and 12 months

This outcome will be analysed using a mixed-effects ordered logistic regression model, with 'course' as a random effect. Site of recruitment, age, gender, HADS depression score, and the outcome at baseline will be included as fixed covariates.

All participants who completed the census global health question score at either 6 or 12 months will be included in the analysis.

MI will be used to account for participants who are missing the outcome at either 6 or 12 months. The MI strategy will be the same as that for the primary and other secondary outcomes, except we will impute the whole score (as there are no individual components).

Total DDDs up to 12 months post-randomisation for psychotropic drugs, drugs for pain, weak opioids, and strong opioids

These outcomes will be analysed using a mixed-effects linear regression model, with ‘course’ as a random effect. Restricted maximum likelihood (REML) will be used. The model will include site of recruitment, age, gender, HADS depression score, and Total DDD in 3 months before randomisation at baseline as covariates. All participants who have data on Total DDD up to 12 months post-randomisation will be included in the analysis. Mean imputation will be used for missing baseline covariates.

Proportion of participants using weak opioids and strong opioids at 12 months post-randomisation

These outcomes will be analysed using a mixed-effects logistic regression model, with ‘course’ as a random effect. The model will include site of recruitment, age, gender, HADS depression score, and weak or strong (depending on outcome) opioid use at baseline (defined as a prescription for weak or strong) opioids in the 12 weeks before randomization) as covariates. All participants who have data on whether they had had a weak/strong opioid prescription at 12 months will be included in the analysis.

7.5 Adherence-adjusted analysis

As a secondary analysis, CPG disability, CPG pain intensity, PSEQ, HADS anxiety, HADS depression, CPAQ, HEIQ, and EQ-5D, all at 12 months will be re-analysed to obtain a complier average causal effect of treatment (CACE). We define ‘compliers’ as those who attend more than half of the course (i.e. those present for at least 12 of the 24 course components). The compliers can only be observed in the intervention arm, where an indicator variable will identify whether the individual complied. The compliers’ class is unobserved in the control arm.

We assume the Stable Unit Treatment Value Assumption (SUTVA), namely: (a) no interference between study units (the outcome for each participant depends only on their own treatment assignment and not the treatment assignment of any other participant), and (b) consistency, which implies that the observed outcome for each participant will equal one of the potential outcomes, no matter how the treatment was received.

In addition for identification, we assume (a) monotonicity: there are no defiers; and (b) exclusion restriction: treatment allocation only has an effect on outcome through treatment received and the effect of assignment is completely mediated by treatment exposure. .

Under the assumptions stated above, we will use randomisation as an instrumental variable for treatment received and obtain a CACE treatment estimate by a two-stage least square instrumental variable regression (using STATA command `ivregress`). We will run two analyses, one without any covariates and another one which includes all the baseline covariates included in the primary analysis models, namely CPG disability score at baseline, site of recruitment, age, gender, and the HADS depression score at baseline. The covariate-adjusted CACE will be considered the primary CACE analysis.

We will assume that missing data are missing at random and use the same multiply imputed datasets produced for the primary analyses. We will analyse each of multiply imputed sets, using robust estimation for the variance (using the option `vce(cluster clustvar)`) to account for the possible clustering by course group; finally obtaining MI estimates using Rubin's rules as before.

7.6 Mediator analyses

We will perform a mediator analysis to obtain the direct and indirect effects of treatment on the CPG disability score at both 6 and 12 months, using self-efficacy (PSEQ) at 12 weeks as a mediator.

We will use a structural linear mean model that allows for the interaction of randomisation with moderator, and perform an instrumental variable analysis (`ivregress` in STATA), using the interaction of randomization and baseline PSEQ as an instrument for the mediator, and including the interaction between randomisation and PSEQ at 12 weeks in the model.

To study the combined effect of compliance and self-efficacy, we will do a second mediation analysis. Let Y denote the outcome (CPG disability score), R the group as randomised, C the binary compliance (as defined in Section 7.5) and S the self-efficacy measure (PSEQ, the mediator). We will use the following structural model:

$E[Y_i(R=1) - Y_i(R=0) | C_i = 1 \ \& \ S_i = s] = \beta_c c + \beta_s s + \beta_{cs} cs$, where β_{cs} represents the effect moderation of self-efficacy on those that comply.

This equation implies an exclusion restriction – the expected treatment effect being zero when less than half of the sessions are attended (though we allow for a self-efficacy to have a non-zero effect on outcome). For identification, we will use randomisation as an instrument for compliance, and randomisation by PSEQ at baseline interaction as an instrument for the mediator.

We will test the strength of the instruments using `estat firststage` post-estimation command in STATA. Low values of the R^2 or F statistic of the joint correlation of the mediator and the two instruments are indicative of weak instruments (rule of thumb F statistic less than 10 indicates weak instruments, Stock and Yogo 2005). If the instruments are weak, the estimates will still be unbiased but the standard error obtained by 2SLS are incorrect; in this case, we will use LIML estimation.

For both instrumental variable regressions, we will use the same multiply imputed datasets as the primary analyses and analyse each of them using the robust standard

error estimate (`vce(cluster clustvar)`) to account for possible clustering by session groups; finally obtaining MI estimates using Rubin's rules as before.

As a sensitivity analysis to our instrumental variables approach, assuming that there is no unmeasured mediator-outcome confounding, we will use the same structural mean model as above on the complete cases, and fit the model with the `paramed` command in STATA which allows for treatment-mediator interactions. We will include CPG disability score, HADS score, HEIQ, CPAQ and EQ5D at baseline in the model as they are considered to be a priori mediator-outcome confounders (by randomisation, there is no confounders of treatment-outcome, and treatment-mediator associations).

For the model estimating the combined effect of compliance and mediator, we will assume we measured all confounders of the mediator-outcome and compliance-outcome associations, these are CPG disability score, HADS score, HEIQ, CPAQ and EQ5D at baseline and include them in the model, which we will fit to the complete case dataset using the command `paramed` in STATA.

7.7 Additional data summaries

The following additional data summaries will be produced:

- The mean (SD) for the change from baseline for CPG disability, CPG pain intensity, PSEQ, HADS anxiety, HADS depression, CPAQ, HEIQ, and the EQ-5D at both 6 and 12 months.
- The effect size (based on Cohen's D, i.e. the treatment effect divided by the standard deviation) for CPG disability, CPG pain intensity, PSEQ, HADS anxiety, HADS depression, CPAQ, HEIQ, and the EQ-5D at both 6 and 12 months.

8. Tables

The following tables will be produced:

Table 1 – baseline characteristics

	Intervention (n=...)	Control (n=...)
Age (years) – mean (SD)		
Male – no. (%)		
Living arrangements – no. (%)		
Alone		
With others		
Ethnicity – no. (%)		
White		
Black		
Asian		
Mixed		
Other		
English language fluency – no. (%)		
Fluent		
Good		
Below average		
Poor		
Age at which formal education ended – no. (%)		
No formal education received		
12 years or less		
13 to 16 years		
17 to 19 years		
20 years or later		
Still in full time education		

Other		
Employment status – no. (%)		
Employed, including self employed (full or part time)		
Unemployed and looking for work		
At school or in full time education		
Unable to work due to long term sickness		
Looking after home/family		
Retired from paid work		
Other		
Time kept from usual activities due to pain in past 6 months		
0-6 days		
7-14 days		
15-30 days		
31 or more days		
State of health – no. (%)		
Very good		
Good		
Fair		
Bad		
Very Bad		
Duration of pain – no. (%)		
0-3 months		
4-12 months		
13 months – 2 years		
3-4 years		
5-6 years		
7-10 years		
More than 10 years		
CPG overall – mean (SD)		

CPG disability – mean (SD)		
CPG pain intensity – mean (SD)		
PSEQ – mean (SD)		
HADS depression – mean (SD)		
HADS anxiety – mean (SD)		
CPAQ – mean (SD)		
HEIQ – mean (SD)		
EQ-5D – mean (SD)		
Number of co-morbidities – median (IQR)		
Total amount of drugs taken above the Defined Daily Dose (DDD) in three months prior to randomisation		
Psychotropic – median (IQR)		
Weak opioids – median (IQR)		
Strong opioids – median (IQR)		
Analgesics (including opioids, non-opioids, NSAIDS and other CNS drugs, and oral and topical preparations)– median (IQR)		
Drugs taken orally for neuropathic pain – median (IQR)		
NSAID analgesics (both oral and topical) – median (IQR)		
Proportion of participants prescribed weak opioids – no. (%)		
Proportion of participants prescribed strong opioids – no. (%)		

Table 2 – Number (%) of participants included in each analysis

	Intervention (n=...)	Control (n=...)
CPG disability		
CPG pain intensity		
PSEQ score		
HADS Anxiety score		
HADS Depression score		
CPAQ score		
HEIQ score		
EQ-5D		
Census global health question		
Total amount of drugs taken above the Defined Daily Dose (DDD) in up to 12 months post-randomisation		
Psychotropic		
Weak Opioids		
Strong Opioids		
Analgesics (including opioids and other CNS drugs)		
Proportion of participants using opioids at 12 months post-randomisation		
Weak opioids		
Strong opioids		

Table 3 – Main results for primary and secondary outcomes

	Intervention (n=...)	Control (n=...)	Treatment effect (95% CI)	P-value
CPG disability – mean (SD)				
12 months				
6 months				
CPG pain intensity – mean (SD)				
12 months				
6 months				
PSEQ score – mean (SD)				
12 months				
6 months				
HADS Anxiety score – mean (SD)				
12 months				
6 months				
HADS Depression score – mean (SD)				
12 months				
6 months				
CPAQ score – mean (SD)				
12 months				
6 months				
HEIQ score – mean (SD)				
12 months				
6 months				
EQ-5D – mean (SD)				
12 months				
6 months				
Census global health question – mean (SD)				

12 months				
6 months				
Total amount of drugs taken above the Defined Daily Dose (DDD) in up to 12 months post-randomisation – median (IQR)				
Psychotropic				
Weak opioids				
Strong opioids				
Analgesics (including opioids and other CNS drugs)				
Proportion of participants using opioids at 12 months post-randomisation – no. (%)				
Weak opioids				
Strong opioids				

Table 4 – Results from sensitivity analyses for primary outcome

	Treatment effect (95% CI)	P-value
Main analysis		
Complete case analysis		
Multivariate analysis		
Different imputation model		
CACE analysis		
Re-definition of primary outcome		

Table 5 –Subgroup analyses for primary outcome (CPG disability at 12 months)

Subgroup	Intervention – mean (SD)	Control – mean (SD)	Treatment effect (95% CI)	P-value for interaction
Non-pain				
Co-morbidity				
0-3 (n=...)				
4 or more (n=...)				
Living arrangements				
Living alone (n=...)				
Living with others (n=...)				
PSEQ				
0-20 (n=...)				
21-39 (n=...)				
40-60 (n=...)				
Socioeconomic status				
Lower (n=...)				
Higher (n=...)				
Pain related				
Pain duration				
0-12 months (n=...)				
13 months to 4 years (n=...)				
5 or more years (n=...)				
CPG intensity				
0-3 (n=...)				
4-7 (n=...)				
8-10 (n=...)				
CPG disability				
0-3 (n=...)				

4-7 (n=...)				
8-10 (n=...)				
HADS depression score				
0-10 (n=...)				
11-21 (n=...)				

Table 6 – Courses and activities outside of COPERS during follow-up period

	Intervention (n=...)	Control (n=...)
Courses or activities attended during follow-up period outside of the COPERS trial		
Pain management – no. (%)		
Expert participant programme or other self-management course – no. (%)		
Other wellness or wellbeing courses – no. (%)		
Return to work courses – no. (%)		
Received psychological counseling or therapies – no. (%)		
Frequency of practicing relaxation and/or meditation during follow-up period – no. (%)		
Daily		
Weekly		
Monthly		
Rarely		
Never		

Table 7 – Change from baseline summaries

Outcome	Change from baseline – mean (SD)	
	6 months	12 months
CPG disability		
CPG pain intensity		
PSEQ score		
HADS Anxiety score		
HADS Depression score		
CPAQ score		
HEIQ score		

EQ-5D		
-------	--	--

Table 8 – Standardised differences based on Cohen’s D

Outcome	Treatment effect (95% CI)
CPG disability	
12 months	
6 months	
CPG pain intensity	
12 months	
6 months	
PSEQ score	
12 months	
6 months	
HADS Anxiety score	
12 months	
6 months	
HADS Depression score	
12 months	
6 months	
CPAQ score	
12 months	
6 months	
HEIQ score	
12 months	
6 months	
EQ-5D	
12 months	
6 months	

*Effect sizes were calculated by dividing the treatment effect and the confidence limits by the estimated standard deviation

Table 9 – Differences between responders and participants lost to follow-up

	Responder (n=...)	Lost to follow-up (n=...)	Odds ratio for non-response (95% CI)	P-value
Age (years) – mean (SD)				
Male – no. (%)				
Ethnicity – no. (%)				
White				
Black				
Asian				
Mixed or other				
English language fluency – no. (%)				
Fluent or good				
Below average or poor				
Age at which formal education ended – no. (%)				
Employment status – no. (%)				
Employed				
Other				
CPG disability at baseline – mean (SD)				
CPG pain intensity at baseline – mean (SD)				
PSEQ at baseline – mean (SD)				
HADS depression at baseline – mean (SD)				
HADS anxiety at baseline – mean (SD)				
CPAQ at baseline – mean				

(SD)				
Number of co-morbidities – median (IQR)				

Table 10 – ICC estimates

Outcome	ICC
CPG disability	
12 months	
6 months	
CPG pain intensity	
12 months	
6 months	
PSEQ score	
12 months	
6 months	
HADS Anxiety score	
12 months	
6 months	
HADS Depression score	
12 months	
6 months	
CPAQ score	
12 months	
6 months	
HEIQ score	
12 months	
6 months	
EQ-5D	
12 months	
6 months	
Census global health question	
12 months	
6 months	
Total amount of drugs taken above the Defined Daily Dose (DDD) up to 12 months post-randomisation	
Psychotropic	
Weak opioids	

Strong opioids	
Analgesics	
Proportion of participants using opioids at 12 months post-randomisation	
Weak opioids	
Strong opioids	
Compliance	
Continuous scale (0-24 components attended)	
Binary scale (attended 12 or more components)	

Appendix I. Methods of calculating derived variables

CPG disability at 6 months

This is derived in the same method as the CPG disability score at 12 months (as described in section 2.3).

CPG pain intensity score at 6 and 12 months.

This is a composite of three questions which assess the participant's pain intensity at present, and the maximum and average intensity over the past 6 months. Each question is scored on a scale of 0-10. The outcome is the mean of the three questions, multiplied by 10. Its range is from 0-100, with higher scores indicating worse pain.

PSEQ (Pain Self-Efficacy Questionnaire) score at 6 and 12 months

This is a composite of 10 questions which ascertain the participant's level of confidence to live a normal life despite their pain. Each question is scored on a scale of 0-6. The outcome is the sum of all 10 questions. Its range is 0-60, with higher scores indicating higher levels of confidence.

HADS (Hospital Anxiety and Depression Scale) Anxiety score at 6 and 12 months

This is a composite of 7 questions which ascertains the extent of the participant's anxiety (these are the odd number questions of the HADS questionnaire). Each question has four answers ranging from not experiencing a symptom at all scored as 0, to experiencing a symptom nearly all the time scored as 3. The outcome is the sum of each question. Its range is 0-21, with higher scores indicating more severe anxiety.

HADS (Hospital Anxiety and Depression Scale) Depression score at 6 and 12 months

This is a composite of 7 questions which ascertains the extent of the participant's depression (these are the even number questions of the HADS questionnaire). Each question has four answers ranging from not experiencing a symptom at all scored as 0, to experiencing a symptom nearly all the time scored as 3. The outcome is the sum

of each question. Its range is 0-21, with higher scores indicating more severe depression.

CPAQ (Coping Pain and Acceptance Questionnaire) score at 6 and 12 months

This is a composite of 20 questions which ascertain the participant's ability to cope with their pain. Each question is scored on a scale of 0-6, with 0 indicating the statement is never true, and 6 indicating the statement is always true. There are two subscales: Pain Willingness and Activities Engagement. The statements in the Pain Willingness subscale are reverse scored, so that an answer of 'Always true' gives a score of 0, and a score of 'Never true' gives a score of 6. The outcome is the sum of each question. Its range is 0-120, with higher scores indicating a better ability to cope.

HEIQ (Health Education Impact Questionnaire) score at 6 and 12 months

This is a composite of 5 questions which ascertain the extent to which the participant is able to enjoy life. Each question has four answers ranging from Strongly Agree (scored as 4) to Strongly Disagree (scored as 1). The outcome is the sum of each question. Its range is 4-20, with higher scores indicating more enjoyment in life.

EQ-5D at 6 and 12 months

This is a composite of 5 questions which ascertain whether the participant has any problems with mobility, self-care, performing their usual activities, pain or discomfort, or anxiety or depression. Each question has three answers ranging from 'No problems' (scored as 1) to the worst category (scored as 3). The outcome score will be derived using the method described in the SPSS manual.

CPG overall (baseline variable)

The CPG overall score is a composite of the CPG disability, the CPG pain intensity, and another question assessing the number of days off usual activities due to pain. This question has four categories: 0-6 days, 7-14 days, 15-30 days and 31 or more days. Categories are assigned 0 points for 0-6 days through to 3 points for 31 + days.

CPG pain intensity is grouped as <50 vs ≥ 50 , and CPG disability is grouped as 0 (0-29 points), 1 (30-49 points), 2 (50-69 points), or 3 (70-100 points). An overall

disability score is then formed by adding the points from the grouped CPG disability score (range 0-3) to the points assigned for the number of days off work (range 0-3), giving an overall range of 0-6.

CPG Calculation

Grade 0 Pain free: No pain problems in the last 6 months

Grade I Low pain disability and low pain intensity: Characteristic pain intensity <50 and <3 disability points

Grade II Low disability-high intensity: Pain intensity of 50 or more and <3 disability points.

Grade III High disability- moderately limiting: 3-4 disability points, regardless of pain intensity

Grade IV High disability – severely limiting: 5-6 disability points

Drugs Data Analysis

Total Defined Daily Doses (Total DDD) consumed

The Total DDD for each drug is defined as:

$$\text{Total DDD}_{\text{DrugA}} = (\text{Strength}_{\text{MedA}} \times \text{quantity}_{\text{MedA}}) / \text{DDD}_{\text{MedA}}$$

The Total DDD for a group of medications (e.g. the Total DDD for opioids) is the sum of the Total DDD for each drug within that medication group (e.g. each drug which is considered an opioid). For example, if there are three drugs (drugs A, B, and C), the $\text{TotalDDD}_{\text{opioid}}$ is defined as:

$$\text{TotalDDD}_{\text{opioid}} = \text{TotalDDD}_{\text{DrugA}} + \text{TotalDDD}_{\text{DrugB}} + \text{TotalDDD}_{\text{DrugC}}$$

The DDD (used in the denominator of the calculation for the TotalDDD) is determined in the first instance by the WHO register, then by precedent in other trials (OPERA and TOIB), and then by clinician consensus. For compound drugs, e.g. co-codamol we will separate out components (paracetamol & codeine) and work out the DDD for each component drug.

Data

Medications used over a 15 month period have been collected from GP participant records. We extracted drug name and strength used, plus quantity and the dates i.e. number of times the medication was prescribed. We have used the prescription cost analysis database to attach a cost to each individual preparation used. Using the World Health Organization (WHO)-defined daily dose for each drug we will generate number of days of medication used by *British National Formulary* chapter and subchapter.

Outcomes

We consider the following outcomes:

- 1) Total Defined Daily Doses (Total DDD) consumed of psychotropic drugs (Table 11) up to 12 months post randomisation
- 2) Total DDD consumed of all analgesics up to 12 months post randomisation

- 3) Total DDD consumed of weak opioids up to 12 months post-randomisation (as defined by BNF 4.7.2 are codeine, dihydrocodeine and meptazonol)
- 4) Total DDD consumed of all NSAID analgesics (oral and topical combined) up to 12 months post randomisation
- 5) Total DDD consumed of all CNS drugs for neuropathic pain (see Table11) up to 12 months post-randomisation
- 6) Total DDD consumed of strong opioids up to 12 months post-randomisation (as defined by BNF 4.7.2, all opioids prescribed other than the ones listed above as weak)

Calculations for psychotropic drugs will be based on BNF subchapters 4.1, and 4.3, opioids based on BNF paragraph 4.7.2, and analgesics including opioids based on BNF paragraphs 4.7.1, 4.7.2, 4.7.3, and paragraphs 10.1.1, 10.2.2, and 10.3.2.

Relevant Drugs

We will work out DDD for BNF chapter 4 and 10 groups of drugs, these are drugs used for treating chronic pain (see table below). We will exclude all drugs administered as injections, but we will include soluble drugs, gels and liquids.

Table 11- Pain related drugs

	Chapter	Subchapter	Paragraph	Comments
Psychotropic drugs	4. Central Nervous System	4.1. Hypnotics and Anxiolytics	4.1.1 Hypnotics 4.1.2. Anxiolytics	NOT: chloral and derivatives, clomethiazole or antihistamines
		4.3. Antidepressant drugs	4.3.2 Monoamine-oxidase inhibitors 4.3.3. Selective serotonin re-uptake inhibitors 4.3.4 Other antidepressant drugs	

Analgesic drugs		4.7 Analgesics	4.7.1 Non opioid analgesics 4.7.2. Opioid analgesics 4.7.3 Neuropathic and functional pain	4.8.1 Gabapentin and pregabalin feature as an anti-epileptic but also feature in 4.7.3 Neuropathic and functional pain For this analysis 4.3.1 tricyclic anti-depressants are included in section 4.7.3
	10. Musculoskeletal and joint diseases (exclude steroids, DMARDS)	10.1 Drugs used in rheumatic diseases and gout	10.1.1 Non-steroidal anti inflammatories	Exclude aspirin No steroids
		10.2 Drugs used in neuromuscular disorders	10.2.2 Skeletal muscle relaxants	
		10.3 Drugs for the relief of soft tissue inflammation	10.3.2 Rubefaciants and other topical anti-rheumatics	Not enzymes

REFERENCES

1. Foster G, Taylor SJC, Eldridge SE, Ramsay J, Griffiths CJ: Self-management education programmes by lay leaders for people with chronic conditions. Cochrane Database of Systematic Reviews 2007(4):CD005108.
2. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, Bouter LM, de Vet HC: Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine 2008, 33(1):90-94.
3. StataCorp: Stata Statistical Software. Release 10.1. In.: College Station, TX StataCorp LP; 2007.
4. Parsons S, Breen A, Foster NE, Letley L, Pincus T, Vogel S and Underwood M. Prevalence and comparative troublesomeness by age of musculoskeletal pain in different body locations. Family Practice 2007; 24: 308–316.
5. Nicholas NK. The pain self-efficacy questionnaire: taking pain into account. Eur J pain 2007; 11: 153-163
6. Nicholas NK, Asghari A, Blyth F. What do the numbers mean? Normative data in chronic pain measures. Pain. 2008; 134: 158–173