
# CASPER

## Collaborative care and active surveillance for screen-positive elders with sub-clinical depression: a pilot study and definitive randomised evaluation (ISRCTN 0220295)

STATISTICAL ANALYSIS PLAN

Clinical Effectivenss Analysis of the CASPER Trial

Version 2.1

**York Trials Unit**
Department of Health Sciences
University of York
York, YO10 5DD

**Version date:** 04/09/2014
**Author:** Ada Keding
**Chief Investigator:** Prof Simon Gilbody
**Trial Coordinator:** Dr Helen Lewis
**Trial Statistician:** Ada Keding
**Second Statistician:** Dr Catherine Hewitt

# Contents

# 1. Glossary

| | | |
|---|---|---|
| BA | Behavioural Activation | |
| CASPER | Collaborative Care in Screen-Positive Elders | |
| | CASPER Study | Study designed to assemble an epidemiological cohort of people over 65 years of age |
| | CASPER Cohort | Cohort of participants over 65 years of age who consented to be followed up as part of the CASPER Cohort |
| | CASPER Trial | Trial using CASPER Cohort participants with *sub-threshold* depression who consented to take part in the CASPER trial |
| | CASPER PLUS Trial | Trial using CASPER Cohort participants with *above-threshold* depression who consented to take part in the CASPER PLUS trial |
| CC | Collaborative Care | |
| CONSORT | Consolidated Standards of Reporting Trials | |
| CRF | Case Report Form | |
| IAPT | Improving Access to Psychological Therapies | |
| ITT | Intention-to-treat | |
| MINI | Mini-International Neuropsychiatric Interview (Major Depressive Episode Module used in this study) | |
| NICE | National Institute for Health and Care Excellence | |
| NSAE | Non-Serious Adverse Event | |
| PC-MIS | Patient Case Management Information System | |
| SAE | Serious Adverse Event | |
| SAP | Statistical Analysis Plan | |
| SOP | Standard Operating Procedure | |

# 2. Introduction

## The CASPER Study

The CASPER study has been designed to assemble an epidemiological cohort of people over 65 years of age (the CASPER Cohort), from which those eligible to participate in a trial of collaborative care for sub-threshold depression (the CASPER Trial including an internal pilot) and a trial of collaborative care for above-threshold depression (the CASPER PLUS Trial) were identified.

## Research Objectives

CASPER Pilot Trial

1. To develop a low intensity collaborative care intervention based upon evidence-supported models of care for elderly people with screen-positive sub-threshold depression
2. To establish the acceptability and uptake of this service by elderly people with screen-positive sub-threshold depression in primary and residential care settings
3. To test the feasibility of conducting a successful trial of low intensity collaborative care intervention for elderly people with screen-positive sub-threshold depression
4. To validate the Whooley questions in a UK elderly population

CASPER Trial

5. To establish the clinical effectiveness of a low intensity collaborative care intervention for elderly people with screen-positive sub-threshold depression
6. To examine the cost effectiveness of a low intensity collaborative care intervention for elderly people with screen-positive sub-threshold depression across a range of health and social care costs

CASPER PLUS Trial

7. To establish the clinical effectiveness of a low intensity collaborative care intervention for elderly people with screen positive above-threshold depression
8. To examine the cost effectiveness of a low intensity collaborative care intervention for elderly people with screen positive above-threshold depression

## Scope of Statistical Analysis Plan

This analysis plan exclusively covers details of the statistical analysis of treatment efficacy in the CASPER Trial: To establish the clinical effectiveness of low intensity collaborative care intervention for elderly people with screen-positive sub-threshold depression (Research Objective 5). Any analyses addressing the trial pilot objectives, the qualitative evaluation and economic analysis as well as any further planned analyses of data from CASPER Cohort participants including the CASPER PLUS trial are detailed elsewhere.

# 3. Trial Design

## Summary

The CASPER trial is a 2-arm, unblinded, pragmatic randomised controlled trial with randomisation at the patient level. The two arms are:

(1) Collaborative Care with Behavioural Activation (CC)

- Customised low intensity care programme delivered to an established protocol by a case manager (mental health / IAPT worker)
- 8-10 weekly sessions, first session conducted face to face, subsequent sessions fact to face or by telephone according to patient preference
- Elements of CC: telephone support, symptom monitoring and active surveillance, low intensity psycho-social management / behavioural activation

(2) Usual Primary Care (UC)

- GP delivered usual care management of sub-threshold depression in line with NICE guidance

Full details of the background and design of the trial are given in the study protocol (latest version at time of writing: Version 2.8, 3 Dec 2013).

## Sample Size

To detect a minimum standard effect size of 0.3 with 80% power and a two-sided 5% significance level would require 352 patients (176 in each group). Although this is an individually-randomised trial, there may be potential clustering at the level of each collaborative care case manager and hence the sample size was inflated to account for this. Based upon an ICC = 0.02 and a caseload size of 20, the design effect would be 1.38 (1 + (20 - 1) × 0.02) and 486 patients (243 in each group) would be required.

The original trial sample size allowed for 10% loss to follow-up, resulting in a total of 540 patients (270 in each group). Following the inspection of interim recruitment figures and disproportionate dropout from the collaborative care arm, this assumption was revised (Substantial Amendment 17). Allowing for a loss to follow-up of 26%, the final sample size needed is 658 patients (329 in each group).

## Randomisation

Participants with sub-threshold depression who met the inclusion criteria, provided written consent and were not actively receiving psychological therapy were eligible for randomisation into the trial. Once invitation, baseline and diagnostic interview data were collected and entered into the study database, randomisation was carried out by the York Trials Unit Randomisation Service. Participants were randomised on a 1:1 basis using simple randomisation to either the intervention group (collaborative care) or control group (usual care).

### Blinding

Blinding of the participants was not feasible, nor was blinding of all members of the study team who were actively involved in the administration of the study. However, the trial statistician responsible for the final statistical analysis will be kept blind to group allocation until the primary analysis has been completed and checked by the second statistician. Some of the trial data explicitly reveal treatment allocation (e.g. intervention details and patient comments), therefore analysis and reporting following the primary analysis will be unblinded.

### Follow-up

Follow-up was at 4 and 12 months post-randomisation by postal questionnaire for all participants who consented to participate in the trial. If non-returned, a reminder letter and second questionnaire pack were sent 21 days following the original mailing date. If the questionnaire was not returned after 35 days, some participants were called in person.

## 4. Outcomes

### Primary Outcome

The primary outcome of the CASPER trial is Depression Severity and Symptomatology (PHQ-9) at 4 months post-randomisation.

*PHQ-9*

The PHQ-9 is a nine item depression scale. Each item is scored between 0 and 3, thus PHQ-9 scores can range from 0 to 27 with higher scores indicating greater depression. Up to two missing questionnaire items are pro-rated, and no total scores are calculated for responses with more than two missing items. Scores of 5, 10, 15, 20 have been used as cut points for mild, moderate, moderately severe and severe depression.

### Secondary Outcomes

- Depression severity and symptomatology at 12 months (PHQ-9)
- Binary Depression Severity at 4 & 12 months (PHQ-9), using scores ≥ 10 to designate moderate depression caseness
- Quality of Life at 4 & 12 months (SF-12 and EQ-5D)
- Psychological Anxiety at 4 & 12 months (GAD-7)
- Mental Health Medication at 4 & 12 months
- Mortality at 4 & 12 months

*SF-12*

The SF-12 is a generic health status measure and a short form of the SF-36 health survey. It consists of 12 questions measuring 8 domains (Physical, Role Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role Emotional and Mental Health) rated over the past month. Questions have 3 or 5 response categories, and responses are summarised into a physical and mental component

score (PCS and MCS). Outcomes range from 0 (lowest level of health) to 100 (highest level of health).

*EQ-5D*

The EQ-5D is a standardised measure of current health status developed by the EuroQol Group for clinical and economic appraisal. The EQ-5D consists of five questions, each assessing a different quality of life dimension (Mobility, Self-care, Usual activities, Pain/Discomfort and Anxiety/Depression). Each dimension is rated on three levels: no problems (score=1), some problems (score=2) and extreme problems (score=3). A weighted summary index can be derived to give a score between 1 (perfect health) an 0 (death). For the purpose of the present analysis, only scores of the individual dimensions will be utilised. The summary index will be analysed separately as part of the cost utility analysis.

*GAD-7*

The GAD-7 is a brief measure of symptoms of anxiety, based on diagnostic criteria described in DSM-IV. It consists of seven questions and is calculated by assigning scores of 0, 1, 2, and 3, to the response categories of "not at all", "several days", "more than half the days," and "nearly every day," respectively. GAD-7 total score for the seven items ranges from 0 to 21. Scores of 5, 10, and 15 represent cut points for mild, moderate, and severe anxiety, respectively.

*Mental Health Medication*

Mediation data was captured by self-report on the follow-up questionnaires. Participants indicated any prescribed medication by selecting from a list of ten anti-depressants. Independent objectively collected medication data from GP records will not be included here and will be incorporated in the economic analysis.

*Mortality*

Mortality was established by flagging all randomised participants to the NHS Information Centre at regular intervals. Members of the research team recorded any identified deaths on the study management database.

## Other Collected Data

- Demographic details at invitation
- Whooley Questions at invitation and baseline
- Questions about physical health problems at baseline
- Questions about falls at baseline
- PHQ-9 at diagnostic interview
- M.I.N.I. (Major depressive episode module) at diagnostic interview
- PHQ-15 at baseline, 4 & 12 months
- CD-RISC 2 at baseline, 4 & 12 months
- Adverse events (as arising up to 12 months follow-up)
- Economic Evaluation at baseline, 4 & 12 months

*Demographics*

Patients completing the background information questionnaire following their invitation provided demographic details regarding age, gender, ethnicity, level of education as well as smoking and alcohol consumption habits.

*Whooley Questions*

The Whooley questions comprise two questionnaire items that form a brief depression screening tool. The questions are: 1. “Over the past month, have you been bothered by feeling down, depressed or hopeless?” and 2. “Over the past month, have you been bothered by having little interest or pleasure in doing things?”. Response options are Yes or No.

*Physical Health Problems*

Participants were asked to indicate whether they suffer from any of a list of ten selected health problems (Diabetes, Osteoporosis, High blood pressure, Rheumatoid arthritis, Osteoarthritis, Stroke, Cancer, Respiratory conditions, Eye conditions, Heart disease) or give details of any other health problems.

*M.I.N.I.*

The 9-item major depressive episode module of the Mini International Neuropsychiatric Interview was used to ascertain the presence or absence of depressive symptoms and categorise participants into non-depressed, sub-threshold depressive symptoms or major depressive episode groups. The nine items comprised: 1) Depressed mood, 2) Loss of interest, 3) Significant weight loss or gain or decrease or increase in appetite, 4) Insomnia or hypersomnia, 5) Psychomotor agitation or retardation, 6) Fatigue or loss of energy, 7) Feelings of worthlessness or excessive or inappropriate guilt, 8) Diminished ability to think or concentrate, or indecisiveness, 9) Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or suicide attempt or a specific plan.

*Falls*

Participants were asked in two questions whether and how often they had fallen is the past 12 months.

*PHQ-15*

The PHQ-15 is a 15-item physical health problems questionnaire. Each health issue is rated as 0 (not bothered), 1 (bothered a little) or 2 (bothered a lot). Items are added to form a scale from 0 to 30, higher scores indicating worse symptom severity. Scores of 5, 10 and 15 have been used as cut-offs for low, medium and high symptom severity. Item 4 of the PHQ-15 (menstrual problems) was not deemed appropriate for the elderly Casper patient population and omitted from all questionnaires. Therefore the total possible PHQ-15 score will be 28.

*CD-RISC 2*

The CD-RISC 2 is a 2-item short form of the full Connor-Davidson Resilience Scale (CD-RISC 25). It is a psychological resilience measure with specific items for bounce-back from adversity and adaptability to change. Agreement with the two items is scored from 0 to 4, resulting in a total score of 0 to 8, where a higher score indicates greater resilience.

*Adverse Events*

Serious Adverse Events (SAEs), i.e. any deaths, unscheduled hospitalisations and any other medically important conditions were identified on an ongoing basis and documented on an SAE form. Outcome status and relatedness to the intervention (as assessed by a member of the trial management group) were recorded.

*Economic Evaluation*

Economic questionnaire items covered participants' health resource use, such as hospital visits and stays as well as use of further NHS and other health support services.

## Data Collection Schedule

Table 1: Data Collection Schedule

| | Invitation | Baseline | Diagnostic Interview | 4 month follow-up | 12 month follow-up |
|---|---|---|---|---|---|
| Consent/Decline | • | | | | |
| Demographic questionnaire | • | | | | |
| Whooley questionnaire | • | • | | | |
| Physical health problems | • | | | | |
| Falls questions | • | | | | |
| PHQ-9 | | •[1] | •[2] | • | • |
| SF-12 | | • | | • | • |
| EQ-5D | | • | | • | • |
| GAD-7 | | • | | • | • |
| PHQ-15 | | • | | • | • |
| CD-RISC 2 | | • | | • | • |
| Medication questionnaire | | • | | • | • |
| M.I.N.I. | | | • | | |
| Economic evaluation | | • | | • | • |
| Objective medication data | | • | | • | • |

[1] PHQ-9 at baseline will be reported in the participant baseline characteristics

[2] PHQ-9 at diagnostic interview will be used as covariate in the analyses

# 5. Data

## Data Sources

- Case Report Forms (CRFs)
    - Decline Form
    - Background Information Form
    - Baseline Questionnaire
    - Four Month Follow-up Questionnaire
    - Twelve Month Follow-up Questionnaire

- Study Management Database
    - Number of Invites Sent
    - Consent and Randomisation Dates
    - Centre, Practice and GP Codes
    - Randomised Allocation Group
    - Case Manager Allocations
    - Questionnaire Return Dates
    - Status Changes (death, dropout from treatment, follow-up or trial)
    - M.I.N.I. Diagnostic Interview Responses and Outcome
    - PHQ-9 at Diagnostic Interview
    - Web Completed Questionnaire Data (baseline, four months, twelve months)

- Patient Case Management Information System (PC-MIS)
    - Collaborative Care Session Dates
    - Collaborative Care Completion Status
    - Collaborative Care Session Type (Phone of Face-to-face)

- Local Investigator Spreadsheets
    - Adverse Event Details

All data will be available in csv format before being imported to the statistical analysis software.

## Data Management and Verification

CRFs will be received in paper format by York Trials Unit and scanned in by the data management team. A copy of the CRFs with the variable names from the database is kept in the Trial Master File. Variable names of online data will match those of paper data.

Comprehensive data validation algorithms have been incorporated into the processing of each CRF, including checks for completeness, internal consistency as well as appropriate data formatting and range checks. The data management team will document any violations of these validation rules. At the end of the trial, a final dataset will be handed over to the trial statistician. The statistician will conduct further data checks including checks for duplicate responses and date chronology across questionnaires. The statistician will generate any derived variables as required. Any decision rules, data changes and assumptions made by the statistician following receipt of the final dataset from

data management will be documented in a Trial Assumptions Form. The statistician will not make any changes to the health economics data and release a copy of this together with cleaned demographics and selected outcomes to the trial health economist.

In the event of multiple questionnaire instances being available in paper and/or online format for a patient at the same follow-up time point, the earliest instance containing a minimum of one completed item of the primary outcome (PHQ-9) will be included in the analysis. The questionnaire dates will be based on the automatically logged completion date for web data and patient completed CRF completion date (if missing: date stamp of CRF return date) for paper data. Individual outcome questionnaires completed online will only be considered if completed on the same day as the PHQ-9.

### Relevant Standard Operating Procedures

Data and documents relevant to the statistician will be kept in a Statistical Master File on the R: drive (secure YTU drive) following the directory structure detailed in the YTU SOP "DS01 Directory structure and version control". Access to this folder will be restricted to the trial statisticians at York Trials Unit (Ada Keding and Catherine Hewitt) and the YTU data management team. Other relevant YTU SOPs or guidance documents that will be followed in the conduct of this trial include: S01 Statistical Considerations; SG02 Statistical Reporting.

## 6. Analysis

### Principles

All analyses will be conducted on intention to treat basis (ITT), including all randomised participants in the groups to which they were allocated. Any post randomisation exclusions will be referred to the independent statistician on the DMEC for approval. Analyses will be conducted in Stata version 13 or later, using 2-sided significance tests at the 5% significance level. Results will be presented with 95% confidence intervals where appropriate. The statistician conducting the analyses will remain blind to treatment group until the primary analysis is finalised and checked by the second statistician.

Any missing data will be reported for all predictor and outcome variables by trial arm and response rates compared. No imputations will be undertaken, and the impact of missing data will be explored as part of secondary analyses of the primary outcome.

### Trial Progression

The flow of participants through the trial will be presented in a CONSORT diagram (see Appendix). Summaries of the numbers of participants invited, assessed for eligibility and randomised will be given, and reasons for exclusion will be listed. Reasons for declining to take part in the study will be listed based on any returned decline forms.

The intervention in the collaborative care arm will be described in terms of the number of participants receiving the treatment, the average number of sessions received, average onset of the intervention since randomisation , duration and how many sessions where conducted face to face or over the phone. Reasons for not receiving any collaborative care will be detailed. The number of

case managers delivering the intervention and the average number of participants per case manager will be given.
Frequencies of dropout (death, dropout from treatment, follow-up or trial) will be presented by trial arm. Reasons for dropout, which are available as free text, will be categorised into common themes and frequencies summarised by trial arm.

## Baseline Data

All participant baseline data (demographics from the background information form, outcome data from the baseline questionnaire, PHQ-9 and MINI responses from the diagnostic interview) will be summarised descriptively by trial arm for all randomised participants and all participants included in the primary analysis. No formal statistical comparisons will be undertaken. Continuous measures will be reported as means, standard deviations, minimum, maximum and interquartile ranges; and categorical data will be reported as frequencies and percentages.

## Analysis Population

The analysis population will include all patients in their randomised groups with available outcome data (for the primary analysis: PHQ-9 at 4 months or 12 months follow-up) as well as complete baseline covariates specified for the analysis.
Data from withdrawn participants were retained up to the date of withdrawal, unless they specifically requested for their details to be removed. Where withdrawal was only from the intervention, follow-up data were continued to be collected. If participants were lost to follow-up due to death or migration, their data were included in the main analysis up to where they have been lost to follow-up.

## Analysis of the Primary Outcome

***Primary Analysis***

Unadjusted descriptives of depression severity (PHQ-9) at all follow-up time points will be presented. A covariance pattern linear mixed effects model will be used to compare collaborative care with usual care on PHQ-9 scores at 4 months. Effects of interest and baseline covariates will be specified as fixed effects, and the correlation of observations within patients over time will be modelled by a covariance structure to describe the random effects. The mixed model will provide increased statistical power by utilising all patients with outcomes for at least one follow-up time point.
The outcome modelled will be PHQ-9 at 4 and 12 months. The model will include as fixed effects: time, treatment group and time-by-treatment interaction, adjusting for PHQ-9 at the diagnostic interview and physical/functional limitations (as measured by the baseline SF-12 physical component score). Different covariance structures for the repeated measurements that are available in the analysis software will be explored, and the most appropriate pattern will be used for the final model based on the model information criteria (AIC). The primary endpoint will be the estimate of the effect of the intervention at 4 months, which will be presented with 95% confidence intervals and associated p-values.
Model assumptions of normality of the standardised residuals and of homogeneity of variance of the standardised residuals against fitted values will be checked. If the model assumptions are in doubt,

the outcome data will be transformed prior to analysis. If the specified mixed model fails to converge with the parameters specified as above, individual regressions at each time point will be undertaken instead.

***Secondary Analyses***

1) To explore the potential clustering within collaborative care case managers, descriptive statistics of the PHQ-9 at 4 months will be presented grouped by case manager in the intervention arm. The intra-cluster correlation coefficient (ICC) will be calculated to assess between-case-manager variability. The primary analysis will be repeated adjusting for the clustering by case manager by including case managers as a random effect.

2) Any baseline predictors of PHQ-9 scores at 4 months (age, gender, GAD-7, PHQ-15, mental health medication use and previous history of depression based on M.I.N.I. responses) will initially be identified by individual logistic regressions using $p<0.10$, controlling for PHQ-9 at randomisation. Any remaining predictors from a combined regression will be included as covariates in the primary analysis model.

3) In order to investigate the impact of missing data, any baseline predictors of non-response at 4 months follow-up will be included as covariates in the primary analysis model. Non-response will be defined as the absence of a valid PHQ-9 score, and predictors will be identified initially by individual logistic regressions followed by a combined regression using $p<0.10$.

## Analysis of Secondary Outcomes

Descriptive statistics will be presented for all secondary outcomes by trial arm at the time points collected and presented graphically where appropriate.

*Depression Severity and Symptomatology at 12 months*

The estimate of the effect of the intervention on PHQ-9 scores at 12 months will be extracted from the primary analysis model and presented with 95% confidence intervals and associated p-value.

*Binary Depression Severity*

A logistic mixed effects model will be used to compare PHQ-9 depression caseness (scores ≥ 10) between treatment arms. The model specification and checking will be similar to the primary analysis, but using depression caseness as the outcome. Any covariates identified for the primary outcome will be included in the model. Odds ratios and 95% confidence intervals will be presented for the effect of the intervention at 4 and 12 months.

*Quality of Life*

The primary analysis will be repeated for the continuous scores of the SF-12 physical and mental component scores, adjusting for baseline values of the outcome in question and any relevant covariates identified for the primary outcome. Estimates of the effect of the intervention will be derived at 4 and 12 months.

Frequencies of responses for each EQ-5D dimension (mobility, self-care, usual activities, pain, anxiety/depression) will be presented descriptively by trial arm. EQ-5D data will be formally analysed as part of the economic evaluation.

*Psychological Anxiety*

The primary analysis will be repeated for the continuous total score of the GAD-7, adjusting for baseline values of the questionnaire and any relevant covariates identified for the primary outcome. Estimates of the effect of the intervention will be derived at 4 and 12 months.

*Mental Health Medication*

A logistic mixed effects model will be used to compare whether any mental health medication was prescribed. The model specification and checking will be similar to the primary analysis, but using any self-reported medication (yes or no) as the outcome. Any relevant covariates identified for the primary outcome will be included in the model. Odds ratios and 95% confidence intervals will be presented for the effect of the intervention at 4 and 12 months.

*Mortality*

The number of deaths occurring during the 12 months trial period will be summarised by trial arm and overall. A chi-square test will be used to compare proportions between trial arms if more than five participants died in each arm.

## Analysis of Other Collected Data

Descriptive statistics will be presented for physical symptom severity (PHQ-15) and psychological resilience (CD-RISC 2) by trial arm at the time points collected.

## Sub-group Analyses

No sub-group analyses will be conducted as part of this analysis.

## Adverse Events

Frequencies of any reported adverse events will be summarised by trial arm and overall. Figures will include a breakdown by type of event and relatedness to the intervention. The details of adverse events will be categorised and reported descriptively. The number and percent of patients experiencing at least one adverse event as well as the average number of adverse events per patient will be presented.

## Departures from the Protocol

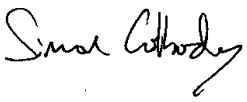
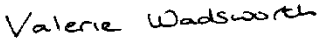
Any details specified in the Casper protocol that have been amended in this analysis plan are detailed in Table 2.

Table 2: Departures from Protocol

| Protocol Version 2.8 (03/12/ 2013) | Statistical Analysis Plan | Reason for Change |
|---|---|---|
| Analyses will be conducted in Stata version 10. | Analyses will be conducted in Stata version 13 or later. | Latest analytical software should be used. |
| The statistician conducting the analyses will remain blind to treatment group until all data summaries and results are finalised. | The statistician will be kept blind to group allocation until the primary analysis has been completed and checked by the second statistician. | Some of the trial data explicitly reveal treatment allocation, and it would not be feasible to ensure blindness throughout. |
| A logistic regression model will be used to compare collaborative care with usual care on the primary outcome. | A covariance pattern linear mixed effects model will be used to compare collaborative care with usual care on PHQ-9 scores at 4 months. | Including the word 'logistic' was an error in the protocol, as the primary PHQ-9 outcome is continuous. In order to maximise the use of available data, mixed models will be used instead of individual regressions at each time point. |
| Appropriate sensitivity analyses will be used to examine the effects of missing data on outcomes. | In order to investigate the impact of missing data, any predictors of non-response will be included as covariates in the primary analysis model only. | As the analysis of secondary outcomes is more exploratory in nature, the sensitivity analysis was restricted to the primary outcome only. |
| Binary PHQ-9 not mentioned. | Binary PHQ-9 (cut-off: ≥ 10 score points) at 4 and 12 months will be included as a secondary outcome. | The published protocol paper lists binary description of the PHQ-9 as a secondary outcome and has been omitted in error from the full protocol. |
| To explore the potential clustering within collaborative care case managers the primary analysis will be repeated adjusting for the clustering using the Huber-White standard estimator (robust standard errors). | The primary analysis will be repeated adjusting for the clustering by case manager by including case managers as a random effect. | As the primary analysis is now a mixed model, a random effect for case managers is appropriate. |

| Protocol Version 2.8 (03/12/ 2013) | Statistical Analysis Plan | Reason for Change |
|---|---|---|
| One of the secondary outcomes: Medication | Medication changed to Mental Health Medication | It was agreed (Casper Data Meeting 25/03/2014) to use only the self-reported questionnaire data in the effectiveness analysis, which refers to mental health medication only. Objective medication data will be analysed as part of the economic analysis. |

# 7. Signatures of Approval

| Name | Trial Role | Signature | Date |
|---|---|---|---|
| Simon Gilbody | Principal Investigator |  | 16th Sept 2014 |
| Helen Lewis | Trial Co-ordinator |  | 18/09/2014 |
| Catherine Hewitt | Senior Statistician |  | 04/09/2014 |
| Ada Keding | Statistician |  | 04/09/2014 |
| Val Wadsworth | Data Manager |  | 04/09/2014 |

# 8. References

Sheehan D, Lecrubier Y, Sheehan H, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. Journal of Clinical Psychiatry. 1998;59 (Suppl 20):22-33.

Ware J, Kosinski M, Keller S. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Medical Care. 1996;34(3):220-33.

EuroQol - a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.

Spitzer RL, Kroenke K, Williams JBW, Lowe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. Arch Intern Med. 2006 May 22, 2006;166(10):1092-7.

Kronke K. The PHQ-9: a new depression diagnostic and severity measure. Psychiatric Annals. 2002;32(9):1-7.

Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: Validity of a New Measure for Evaluating the Severity of Somatic Symptoms. Psychosom Med. 2002 March 1, 2002;64(2):258-66.

Vaishnavi S, Connor K, Davidson J. An abbreviated version of the Connor-Davidson Resilience Scale (CD-RISC), the CD-RISC2: Psychometric properties and applications in psychopharmacological trials. Psychiatry Research. 2007;152(2-3):293-7.

Whooley M, Avins A, Miranda J, Browner W. Case-finding instruments for depression. Two questions are as good as many. Journal of General Internal Medicine. 1997;12(7):439-45.

Kronke K, Spitzer R, Williams J. The PHQ-9: validity of a brief depression severity measure. Journal of General Internal Medicine. 2001;16(9):606-13.

Trials. 2011 Oct 11;12:225. doi: 10.1186/1745-6215-12-225. A randomised evaluation of CollAborative care and active surveillance for Screen-Positive ElderRs with sub-threshold depression (CASPER): study protocol for a randomized controlled trial. Mitchell N1, Hewitt C, Adamson J, Parrott S, Torgerson D, Ekers D, Holmes J, Lester H, McMillan D, Richards D, Spilsbury K, Godfrey C, Gilbody S.

# 9. Appendices

## CONSORT Diagram