OPEN ACCESS

University of BRISTOL

MacNeill, S. J., Wiles, N. J., & Peters, TJ. (2024). *Multi-centre randomised controlled trial of integrated therapist and online CBT for depression in primary care (INTERACT): Statistical Analysis Plan*.

Publisher's PDF, also known as Version of record

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

Multi-centre randomised controlled trial of integrated therapist and online CBT for depression in primary care (INTERACT)

Statistical Analysis Plan

Version 1.0 (24/4/2024)
Based on Protocol version 2.0 (dated 24/3/2023)
ISRCTN 13112900

The following people have reviewed the Statistical Analysis Plan and are in agreement with the contents

	NAME	TITLE	SIGNATURE	DATE
Author	Stephanie MacNeill	Senior lecturer in medical statistics	Stephanie MacNeill	24/4/24
Lead statistician	Prof Tim Peters	Professor of Primary Care Health Services Research	Tim Peters	24/4/24
Study Chief Investigator	Prof Nicola Wiles	Chief investigator	Nicola Wiles	24/4/24
Statistical reviewer	Dr Philip Pallmann	Independent statistician for Trial Steering Committee	Philip Pallmann	24/4/24
Statistical reviewer	Ms Nikki Totton	Independent statistician for Data Monitoring and Ethics Committee	Nikki Totton	24/4/24

Effective Date:	24/4/24

Table of contents

List of abbreviations

Acronym	Details
AE	Adverse Event
AUDIT-PC	Alcohol Use Disorders Identification Test – (Piccinelli) Consumption
BDI-II	Beck Depression Inventory II
BTC	Bristol Trials Centre
CACE	Complier Average Causal Effect
CBT	Cognitive Behavioural Therapy
cCBT	Computerised Cognitive Behavioural Therapy
CI	Confidence Interval
CoBalT	Cognitive behavioural Therapy as an adjunct to pharmacotherapy for primary care patients with treatment resistant depression: a randomised controlled trial
CONSORT	Consolidated Standards of Reporting Trials
DMEC	Data monitoring and ethics committee
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th edition
EQ-5D-5L	EuroQol 5-dimension 5-level
GAD-7	Generalised Anxiety and Depression Assessment
GP	General Practitioner
HEAP	Health Economics Analysis Plan
IAPT	Improving Access to Psychological Therapies
ICD-10	International Classification of Diseases 10th revision
IMP	Investigated Medical Product
INTERACT	Integrated therapist and online CBT for depression in primary care
IQR	Inter-Quartile Range
ITT	Intention to Treat
MICE	Multiple Imputation by Chained Equations
NHS	National Health Service
NICE	National Institute of Health and Care Excellence
NIHR	National Institute for Health and Care Research
OR	Odds ratio
PC-PTSD-5	The Primary Care PTSD Screen for DSM-5
PGfAR	Programme Grants for Applied Research
PHQ-9	Patient Health Questionnaire-9
PTSD	Post Traumatic Stress Disorder
RCT	Randomised Controlled Trial
SAE	Serious Adverse Event (subset of AE)
SAP	Statistical Analysis Plan
SAPAS	Structured Assessment of Personality Abbreviated Scale
SD	Standard Deviation
SSAR	Suspected Serious Adverse Reaction
SUSAR	Suspected Unexpected Serious Adverse Reaction
TMG	Trial Management Group
TSC	Trial Steering Committee
WSAS	Work and Social Adjustment Scale

1. INTRODUCTION AND 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 INTERACT.

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 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.

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 this analysis plan, but if reported the source of such a post-hoc analysis will be declared.

Amendments to the statistical analysis plan will be described and justified in the final report of the trial.

2. BACKGROUND AND RATIONALE

2.1 Rationale

Cognitive Behavioural Therapy (CBT) is an effective treatment for depression, but there is substantial variation in the provision of high intensity treatments across England(1). Computerised CBT interventions (cCBT) were designed to make CBT more accessible and widely available at lower cost. However, adherence to cCBT is often poor and, in the absence of therapist support, effects are modest and short-term(2). Moreover, cCBT is often inflexible and does not allow identification of conditional beliefs or detailed formulations; the latter are crucial elements of CBT(3).

The INTERACT study is a programme of research funded by the National Institute for Health and Care Research Programme Grants for Applied Research (NIHR PGfAR) that aims to integrate online CBT materials and high intensity therapy from an accredited therapist to deliver effective CBT to those who need it. The research programme consists of three stages. The first stage of the research focused on the development of an online platform for delivering integrated high-intensity CBT for depression and gathering design ideas and feedback from key stakeholders such as CBT therapists and patients with experience of CBT. The second stage involved a pilot evaluation of the platform which led to further refinements and training materials for therapists who will deliver the integrated intervention. Thus, this third and final phase of the INTERACT programme is a multi-centre randomised controlled trial (RCT) to fully evaluate the clinical and cost-effectiveness of this integrated approach to CBT for depression. This statistical analysis plan (SAP) relates to the RCT evaluation.

2.2 Trial objectives

The primary objective is to establish the clinical effectiveness and cost-effectiveness of an integrated approach to delivering CBT in reducing depressive symptoms and improving quality of life over 12 months (compared with usual care) for primary care patients with depression. Analyses supporting the cost-effectiveness analysis are described in a separate Health Economics Analysis Plan (HEAP) reported elsewhere.

2.3 Trial design

INTERACT is a pragmatic, two parallel group multi-centre RCT with allocation at the level of the individual.

2.4 Trial centres

The study is based in primary care in three trial centres: University of Bristol (co-ordinating centre), University College London, and Universities of Hull/York. Patients are recruited from primary care general practices (GP practices) in the surrounding areas of Bristol, London and Hull/York.

2.5 Eligibility criteria

2.5.1 Inclusion criteria

Participants are eligible if they:

- Are aged≥ 18 years
- Score ≥14 on the Beck Depression Inventory (BDI-II)
- Meet International Classification of Diseases 10th revision (ICD-10) criteria for depression

2.5.2 Exclusion criteria

Participants are excluded if they:

- Have experienced alcohol or substance dependency in the past year
- Have experienced bipolar disorder
- Have experienced schizophrenia/psychosis
- Have experienced dementia
- Are currently under psychiatric care (including those referred but not yet seen) for depression
- Cannot complete questionnaires unaided or would require an interpreter
- Are currently receiving CBT or other psychotherapy
- Have received high-intensity CBT in the past 4 years
- Are taking part in another interventional trial
- Are not willing or able to receive CBT via computer/laptop/smartphone

2.6 Treatments

2.6.1 Trial intervention – Integrated Cognitive Behavioural Therapy for depression

The intervention comprises of nine therapist-led sessions, with up to a further three sessions if deemed clinically appropriate by the therapist. The first session takes place face-to-face (in person or by videocall) with the subsequent sessions taking place online using the INTERACT platform. Patients in the intervention arm also continue to be cared for by their GP.

It is possible that patients and therapists may reach an “agreed end” of therapy in fewer than nine sessions where clinically appropriate.

2.6.2 Usual care

Participants allocated usual care continue to receive treatment as usual from their GP. This may include referral to local psychological services provided by NHS Talking Therapies services (formerly Improving Access to Psychological Therapies (IAPT)) or antidepressant medication, as appropriate. There are no restrictions on the treatment options than can be offered to this group. However, receipt of psychological therapy received through NHS Talking Therapies services or privately is recorded as part of the follow-up questionnaires (in addition to other data on treatments and health care usage during the trial).

2.7 Recruitment, screening and consent

The trial aims to recruit participants with depression from primary care using record searches and in-consultation recruitment.

2.7.1 Identification of participants

2.7.1.1 Record searches

GP practices conduct searches of their computerised records for potentially eligible patients (defined as those who are aged 18 years or over and have a diagnosis of depression). Practices exclude those who would be unsuitable due to the exclusion criteria. Searches are conducted using a combination of primary care diagnostic codes and manual screening of results lists by practice staff including a practice GP.

Potentially eligible patients are then mailed an invitation to participate by the GP practice asking their permission to be contacted by the research team. One reminder letter will be sent if the patients have not responded after two weeks.

Patients are also able to respond anonymously if they wish to decline participation and are able to provide a reason for declining. Decliners are asked to indicate whether they are willing to be interviewed briefly over the telephone about their reasons for declining and, if so, to add their contact details to the form. Practices are asked to provide anonymised data for all patients identified by the record search (age, gender and reason for exclusion by practice). This information is collected to help report the generalisability of the study results.

2.7.1.2 In consultation

GPs can also identify patients in consultation that they think might be suitable for the trial. They introduce the trial and ask the patient for their permission to be contacted by the research team.

2.7.1.3 Screening

Primary care patients who have been referred (or expressed interest by returning their postal invitation reply form) are telephoned by a researcher from the local site. They briefly explain the study, check that a patient information sheet has been received and answer any questions the participant may have.

The researcher then proceeds with the telephone screening questionnaire which includes questions about: the participant's age; gender; whether they are currently receiving psychotherapy; whether they are receiving care from psychiatric services for their depression; receipt of high-intensity therapy in the last four years; whether they could complete questionnaires unaided; whether they are taking part in another research study which involves receiving an intervention; and whether they would be willing and able to receive CBT online.

If the participant meets the screening criteria they are offered a detailed eligibility screening appointment (baseline assessment) with the researcher.

2.7.1.4 Baseline assessment

At the baseline assessment the local researcher explains the study in more detail, answers questions the participant may have and obtains written informed consent. They check whether the participant's circumstances have changed since they completed the screening questionnaire in order to check if they are still eligible for the baseline assessment.

Potentially eligible participants are then asked to complete a number of questionnaires. These include the Beck Depression Inventory (BDI-II) (4) (a brief measure of depressive symptoms), and the Clinical Interview Schedule – Revised version (CIS-R) (5) (6) an in-depth self-assessment of psychiatric symptoms that establishes whether an ICD-10 diagnosis of depression is met. Sociodemographic questions include: age, gender, employment status, qualifications, ethnicity, housing and marital status. Participants are also asked for relevant medical history including: co-morbidity and history of depression and treatment including antidepressant medication and adherence. They are asked if they are willing for their summary data to be passed to their GP. Participants who score 14 or more on the BDI-II and have an ICD-10 primary diagnosis of depression on the CIS-R are told they are eligible to enter the trial. Eligible participants are asked to complete further questionnaires including: the Patient Health Questionnaire (PHQ-9)(7) and General Anxiety Disorder questionnaire (GAD-7)(8) which are brief measures of depression and anxiety used in psychological services; the EuroQol 5-dimension 5-level (EQ-5D-5L)(9) and the Work and Social Adjustment Scale (WSAS) (10); a simple measure of impairment in functioning. They are also asked to complete the Big Five Personality Inventory neuroticism subscale (11); a personality scale (Structured Assessment of Personality Abbreviated Scale: SAPAS)(12); an alcohol use scale (Alcohol Use Disorders Identification Test – (Piccinelli) Consumption: AUDIT-PC) (13) and a measure of trauma using the Primary Care Post-Traumatic Stress Disorder (PTSD) Screen for Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-5) instrument (PC-PTSD5)(14) . Eligible participants are asked further questions about their history of depression and whether they have ever been referred to a psychiatrist. Additional information is collected on life events, financial stress, social

support and alcohol use(13). Eligible participants are also asked to complete a number of computerised cognitive processing tasks. Analysis of these tasks will be analysed and reported separately.

2.7.1.5 Rescreening

Patients who do not meet the eligibility criteria at telephone screen or baseline may be eligible for rescreening. For example, those who are currently receiving a course of therapy could be rescreened once this course has ended. Patients who do not meet the BDI-II or ICD-10 criteria at baseline can be offered a rescreen as long as at least one month has elapsed since their original baseline appointment.

2.7.1.6 Consent

Prior to commencing the baseline assessment, the researcher obtains written informed consent from the patient relating to their participation in the trial. Informed consent is obtained either via an online consent 'e-consent' method (if the baseline assessment is being conducted remotely) or via paper-based informed consent if the researcher is meeting with the patient face-to-face.

2.8 Randomisation

Randomisation is stratified by centre and minimised on gender, current antidepressant use (Yes/No), and depression severity using BDI-II tertiles derived from the CoBalT study baseline scores (BDI-II ≤25; 26-35; ≥36). If participants indicate a non-binary gender at baseline, we will use an unpredictable computer-generated code to randomly select a binary gender code, prior to randomizing them to a treatment group in the usual way. Details of self-reported gender will be given in full as part of the reporting process.

Stratifying by centre aims to ensure a balance in terms of local differences and also proportionate workload for therapists. The minimisation variables are important prognostic indicators and hence minimising will ensure a balance between the two groups.

2.9 Sample size justification

The primary outcome is the BDI-II score at six months post-randomisation (continuous variable). The National Institute of Clinical Excellence (NICE) depression guidelines group (2004) suggested that 0.35 standard deviations (SDs) represents a clinically important difference, which is approximately 4-5 points on the BDI-II (IPCRESS/CoBalT: standard deviation 12.9/13.9) and the study sample size was estimated on this basis. 173 participants in each group gives 90% power to detect a difference of 0.35 SDs on the BDI-II at a two-sided 5% significance level. Assuming 20% attrition at six months, 434 patients need to be recruited. However, given that referral to IAPT may be part of usual care for the comparator group, it is possible that this may affect the plausibility of the target difference between groups. While it is difficult to determine the impact of this, if the difference to be detected was reduced to 0.30 SDs and a slightly higher follow-up rate (85% in line with previous trial(15)) is achieved, with a total sample size of 434 patients, the study would still have adequate (>80%) power to detect such a difference (see Table below).

	Power	**80% Power**		**90% Power**	
	Attrition at 6 months	**15%**	**20%**	**15%**	**20%**
Difference to be detected	**0.35 SDs**	306	326	408	434
	0.30 SDs	414	440	554	588

The sample size was not inflated to account for clustering by therapist as there was little evidence of any therapist effects within other trials of CBT for depression (15, 16). The intraclass correlation coefficients for the continuous BDI outcome (adjusted for baseline BDI score) were very small for CoBalT (0.0027) and for IPCRESS (the precise value could not be estimated indicating it was almost zero). Hence, inflating the sample size to account for any potential clustering by therapist would be unduly conservative. However, in sensitivity analyses, the methods proposed by Roberts & Roberts (17) will be used to obtain a fully heteroscedastic model to explore any potential therapist effects.

2.10 Blinding

It is not possible to blind participants to their treatment allocation because of the nature of the intervention. To eliminate the potential for observer bias, self-reported outcome measures (e.g. BDI-II) are used to assess outcomes. These outcomes have been widely used in previous depression trials (15, 16, 18).

2.11 Interim analyses

No interim analyses are planned.

2.12 Trial oversight

2.12.1 Trial management group (TMG)

The Trial Management Group (TMG) will be led by the INTERACT programme leads, David Kessler and Nicola Wiles. It will comprise all investigators involved in the trial, the trial manager, research and administrative staff, with input from patient/public representatives. Members of the TMG contribute to the trial in the following ways: trial design, trial centre recruitment and trial conduct, trial management, trial logistics and cost management, CBT expertise, economic evaluation, trials methods, statistical data analysis, and publication. The TMG will meet approximately monthly to oversee the day-to-day management of the trial. The TMG will be provided with detailed information by site staff regarding trial progress. Most meetings will be by teleconference/Skype, but the TMG may be required to meet face to face once or twice a year.

2.12.2 Trial steering committee (TSC)

The TSC is an independent, multidisciplinary group chaired by André Tylee (Professor of Primary Care Mental Health) until July 2021, by Sandra Eldridge (Professor of Biostatistics) until October 2021, and thereafter by Ed Watkins (Professor of Experimental and Applied Clinical Psychology). Other members are Shaun Lawson (Professor of Social Computing), Philip Pallmann (Senior Research Fellow in Statistics), Michael Moore (Professor of Primary Health Care Research), Marta Sawinska (patient representative), and Mark Tucker (patient representative). Their role is to provide independent oversight, progress monitoring, and expert advice during the conduct of the trial. The TSC also includes members of the INTERACT team – David Kessler, Nicola Wiles (Programme Co-Leads), Tim Peters (Statistician) and Debbie Tallon (Programme Manager). Meetings take place at least annually and more frequently if judged necessary. Members of this group have previously acted as the Programme Steering Committee for the INTERACT programme of research.

The TSC will consider and act, as appropriate, upon the recommendations of the Data Monitoring and Ethics Committee and ultimately carries the responsibility for deciding whether a trial needs to be stopped on grounds of safety or efficacy.

2.12.3 Data Monitoring and Ethics Committee (DMEC)

The independent Data Monitoring and Ethics Committee (DMEC) is chaired by Professor David Kingdon (Emeritus Professor of Mental Health Care Delivery) and includes Professor Chris Burton (Professor of Primary Medical Care) and Nikki Totton (clinical trials medical statistician). The DMEC monitors accumulating trial data during the trial and makes recommendations to the TSC as to whether there are any ethical or safety issues that may necessitate a modification to the protocol or closure of the trial. The DMEC convenes prior to TSC meetings and, in accordance with the Trial Terms of Reference for the DMEC, is responsible for assessing safety and efficacy of the trial.

2.13 Outcome measures

2.13.1 Primary outcome

The primary outcome is the BDI-II score at six months post-randomisation, measured as a continuous variable.

2.13.2 Secondary outcomes

The BDI-II score (measured as a continuous variable) at 12 months post-randomisation will be a secondary outcome. The following secondary outcomes are measured at six and 12 months post-randomisation:

- Treatment response (at least 50% reduction in depressive symptoms on the BDI-II compared with baseline)

- Remission of symptoms (BDI-II<10)
- Percentage reduction in depressive symptoms on the BDI-II (i.e. the proportional change)
- Depressive symptoms measured on the PHQ-9
- Quality of life (EQ-5D-5L and WSAS)
- Anxiety (GAD-7)

In addition, the EQ-5D-5L will inform the economic evaluation, as will data on: the number of primary care consultations and prescribed medication collected from practice medical records; the use of other primary and community care services; secondary care related to mental health, private treatments; use of social services; burden on informal care givers; personal costs related to mental health and benefits received. These analyses will be described separately in a Health Economics Analysis Plan (HEAP).

3. GENERAL ANALYSIS CONSIDERATIONS

3.1 Analysis populations

The Full Analysis set includes all randomised participants. A primary intention-to-treat (ITT) analysis will be conducted using this dataset comparing groups as randomised without imputing missing data. Sensitivity analyses will explore the robustness of the primary ITT analysis to assumptions regarding missing data (described in **section 6.3.4**).

Two analyses of safety data will be performed. First, safety analyses will be conducted on all randomised participants according to the group to which they were randomised. Separate to this, these analyses will be performed according to treatment received (19). For the purpose of this analysis, people will be deemed as receiving the intervention if they had at least one INTERACT therapy session.

3.2 Derived variables

The algorithms for the calculation of derived variables (including outcomes and screening variables collected solely at baseline) in this study are described below:

BDI-II Each item of the BDI-II is rated on a 4-pt scale ranging from 0-3 and he BDI-II is scored by summing the ratings for each of the 21 items.

AUDIT-PC The AUDIT-PC scale comprises of 5 items of which eight are scored on a 5-pt scale ranging from 0-4 and one item rated as 0, 2 or 4. The AUDIT-PC score is derived by summing the ratings across each of the items. This variable is collected at baseline only.

CIS-R The CIS-R will be used to identify ICD-10 diagnoses for depression and anxiety, a total CIS-R score across all symptoms and CIS-R depression severity score (summing scores for concentration, fatigue, sleep, depression and depressive ideas). Scores for each symptom range from 0 to 4 (and from 0 to 5 for depressive ideas) where increasing scores reflect higher levels of symptoms. This variable is collected at baseline only.

EQ-5D-5L The EQ-5D-5L questionnaire comprises 5 items each having 5-level responses coded 1-5. NICE currently advises that the 5-level valuation set for England is not recommended for use to derive utilities, instead advising that the validated mapping function to the 3-level valuation set be used (20). We will follow the up-to-date NICE guidance at the time of analysis.

GAD-7 For each of the seven items, scores of 0, 1, 2, and 3 will be allocated to the response categories of “not at all,” “several days,” “more than half the days,” and “nearly every day,” respectively. The total score will be derived by summing the scores for the seven items.

PHQ-9 Each of the nine items in the questionnaire scores symptoms of depression over the last 2 weeks on a scale of 0-3 (0: "Not at all"; 1: "Several days"; 2: "More than half the days"; 3: "Nearly every day"). Scores are summed across items for a total score.

SAPAS Each item of the questionnaire corresponds to a descriptive statement about the person of which seven are scored 0/1 ("No"/"Yes") and one is inversely scored. The overall score is obtained by summing the scores on each of the 8 items.

WSAS Each item enquiries how a person's problems affect their ability to do day-to-day tasks. Items are scored on a scale of 0-8 (0: "Not at all", 2: "Slightly", 4: "Definitely", 6: "Markedly" and 8 "Very severely") and the total score is derived by taking the sum across all items.

For outcomes on the BDI-II, PHQ-9, SAPAS and GAD-7, individual missing items will be addressed using the following rule adopted in the CoBalT study. If >10% of the items are incomplete then the data collected on that measure for that participant will be disregarded. However, if <10% of items on a particular measure are missing, missing item(s) will be imputed using the mean of the remaining items (rounded to an integer). Therefore, when an individual had completed 19 or 20 items for the primary outcome measure (BDI-II) then the remaining one or two items will be imputed. For PHQ-9, SAPAS and GAD-7 the 10% rule will mean that only a single item will be imputed. The number of cases for which values have been imputed will be reported.

3.3 Procedures for missing data

In all tables, missing data will be indicated using footnotes. For the primary outcome of the BDI-II at six months, we will use descriptive statistics to describe the baseline characteristics of patients who do and do not have missing primary outcome data. The impact of missing primary outcome data will also be explored as part of a sensitivity analysis described in **section 6.3.4.**

3.4 Study centre effects

Randomisation of participants is stratified by centre and all analyses will adjust for centre and all minimisation variables.

3.5 Outliers

Prior to analysis the trial statistician will use graphs and descriptive statistics to identify potential outliers in the data. These will be queried with the trial manager who will verify available records to confirm whether or not they are data entry mistakes.

3.6 Visit windows

All questionnaires will be analysed regardless of when they are returned. The median time between baseline and receipt of questionnaires will be presented, however, and sensitivity analyses will be performed to assess whether the primary analysis results are affected by adjustment for the timing of the return of questionnaires.

4. DESCRIPTION OF PARTICIPANT CHARACTERISTICS

4.1 Disposition

The flow of participants through the trial will be summarised in a Consolidated Standards of Reporting Trials (CONSORT) diagram that will include numbers relating to the identification, invitation, eligibility, reasons for exclusion, participants consenting, numbers of participants randomised to the two treatment groups, losses to follow-up and the numbers analysed for the primary outcome.

4.2 Baseline characteristics

Baseline characteristics of participants will be compared between the two arms by reporting relevant summary statistics to determine whether any potentially influential imbalances have occurred (albeit by chance). Baseline characteristics will be summarised using the mean (SD), median (Inter-quartile-range; IQR) or number (%) depending on the nature of the data and its respective distribution. If the baseline characteristics of the groups differ by more than 10 percentage points, or 0.5SDs, then the effect of this variable on the outcome will be investigated in sensitivity analyses.

5. ASSESSMENT OF STUDY QUALITY

5.1 Eligibility checks

The numbers of participants excluded and reasons for exclusions will be described.

5.2 Selection bias

Where patients' age and gender information are available at the screening stage we will use descriptive statistics to compare those who did and did not attend baseline assessment screening (either by actively declining or failing to attend), those who were and were not deemed eligible at the baseline assessment and those who did and did not consent to randomisation.

5.3 Data validation

Once the data are downloaded by the trial statistician, internal consistency checks will be performed in preparing the data for analysis in Stata. These aim to identify spurious values or inconsistencies in responses. When inconsistencies are identified, as with the outliers covered in **section 3.5**, these will be reported to the trial manager who verifies available records.

5.4 Study completion

For the purposes of reporting, we define the end of trial as the collection of the last data item for trial participants. This will be the health care resource use data collected from patients' primary care records once they have completed the 12-month follow-up. Cleaning of the data is an ongoing process and the database will be locked once all data queries are resolved and the last data item is collected. Final analyses will be run once the database is locked. The numbers of patients followed-up and lost to follow-up will be reported for each treatment arm in the CONSORT Flow Diagram.

5.5 Protocol deviations

There will be no prospective, planned deviations or waivers to the protocol. Any protocol breaches will be documented and reported to the Trial Manager, Chief Investigator and Sponsor immediately. Information about protocol breaches will also be included in routine reports to the DMEC and TSC.

6. ANALYSIS OF EFFECTIVENESS

Stata version 18 (or higher) will be used for all INTERACT analyses. Two-tailed tests will be used with effect estimates, 95% confidence intervals (CIs) and p-values presented, with no adjustment for multiple testing. Analyses using regression models will adjust for stratification and minimisation variables as well as baseline values of the outcome involved. The primary approach for analysis will be on an intention-to-treat (ITT) basis defined as analysing participants according to the arm to which they were randomised. A complete case analysis approach will be used and the impact of missing data will be studied in separate sensitivity analyses.

6.1 Mis-randomised patients

Patients will be analysed according to the arm to which they were randomised.

6.2 Summary of primary and secondary endpoints

The primary and secondary endpoints are summarised below:

Outcome	Measure	Timepoints	Interpretation	Range
Primary				
Depression	BDI-II	Baseline and 6 months (also at 12 months)	Higher scores correspond to more severe symptoms of depression	0-63
Secondary				
Treatment response	BDI-II	6 and 12 months	Binary measure where 0 indicates a proportional change in BDI-II from baseline<50%; 1 indicates a proportional change≥50%	0/1
Remission – BDI-II<10	BDI-II	6 and 12 months	Binary measure where 0 indicates that BDI-II scores are ≥10; 1 indicates that scores are <10	0/1
Proportional change in BDI-II from baseline	BDI-II	6 and 12 months	Values further from 0 reflect greater proportional change since baseline. Values >0 indicate an increase in symptoms; values <0 reflect a reduction	-1 to 3.5
Anxiety	GAD-7	Baseline, 6- and 12-months	Higher scores correspond to more severe symptoms of anxiety	0-21
Depressive symptoms	PHQ-9	Baseline, 3-, 6-, 9- and 12-months	Higher scores indicate worse symptoms of depression	0-27
Quality of life	EQ-5D-5L *	Baseline, 3-, 6-, 9- and 12-months	Continuous measure with larger values reflecting better quality of life	-0.594-1
	WSAS		Higher scores indicate greater levels of impairment in ability to undertake day-to-day activities	0-40

* Analysis of the EQ-5D-5L post-baseline as an outcome measurement will be done as part of the health economic analysis.

6.3 Primary analysis

The primary outcome is BDI-II score collected at six months post-randomisation. It will be described in each treatment group using means and SDs. Comparisons between treatment arms will be made using a multivariable linear regression model adjusting for baseline BDI-II scores and variables used in the randomisation by including these as fixed effects in the model. The results will be presented as the adjusted difference between group means, corresponding 95% CI and p-value.

We will perform regression model diagnostics using graphs and summary statistics. Alternative methods of analysis will be considered if the assumptions of the model are not met.

A number of analyses are proposed to assess the sensitivity of the primary analysis to various assumptions. These are described below. Sensitivity analyses will be presented alongside those of the primary analysis so they can be compared and contrasted. As these will be exploratory in nature, 95% CIs and p-values will be presented, but will be interpreted with due caution.

6.3.1 Imbalance between treatment groups

Should there be evidence of imbalance between treatment groups on important baseline characteristics as described in **section 2.7.1.4** and **section 4.2**, sensitivity analyses will be conducted where the primary analysis is repeated, adjusting for variables showing an imbalance. This sensitivity analysis will be performed for the primary outcome.

6.3.2 Complier average causal effect (CACE) analysis

Recognising the inherent bias in estimates derived from per protocol analyses, we will conduct a CACE analysis for the primary outcome. The CACE estimates will be obtained using instrumental variable regression including the same variables used in the primary analysis with randomised group as the instrumental variable and the indicator variable for compliance. Compliance will be based on attendance at an adequate number of therapy sessions defined as attending at least 6 sessions or reaching a jointly agreed end to therapy. Sensitivity analyses will be performed to explore the impact of slight changes to this definition depending on observed patterns in therapy attendance.

6.3.3 Number of therapy sessions attended

Separate to the CACE analysis described above, another sensitivity analysis will explore whether there is a 'dose-response' relationship in terms of the number of therapy sessions attended. Here, the primary analysis model will be re-run with the treatment variable equal to the number of therapy sessions attended – that is, as a continuous variable without a threshold for 'compliance'.

6.3.4 Missing outcome data

The sensitivity of the primary analysis to the impact of missing data will be investigated. The amount of missing data will be compared between arms and descriptive statistics will be used to explore whether there are variables associated with missingness. Where available, reasons for missingness will be reported.

A number of approaches to missing data will be considered then compared and contrasted with the primary analysis. These include imputing missing outcome data using reasonable assumptions for "better" and "worse" case scenarios and multiple imputation by chained equations (MICE). The latter will incorporate analyses to ascertain the sensitivity to the assumption of missingness-at-random.

6.3.5 Therapist effect

As outlined in **section 2.9**, the sample size calculation did not account for clustering by therapist as there was little evidence of therapist effects across other trials of CBT. To explore the potential for therapist effects, a separate analysis using generalised linear and latent mixed models will be used to obtain a fully specified heteroscedastic model following the methods of Roberts & Roberts (18).

6.3.6 Timing of the return of questionnaires

As outlined in **section 3.6**, the primary analysis will incorporate all questionnaires returned regardless of whether they are returned on time or not. Descriptive statistics will be used to describe the timing of the return of questionnaires and the data will be explored to determine whether this differed by treatment group or by BDI-II scores. A sensitivity analysis of the primary outcome will additionally adjust for the timing of the return of these questionnaires in order to assess the impact of late returns.

6.4 Secondary outcomes analyses

The effect of the intervention on the secondary outcomes collected at six and 12-months post-randomisation will be examined using linear regression for continuous outcomes and logistic regression for binary outcomes adjusted for baseline values of the outcome being investigated and variables used in the randomisation. Results will be presented as adjusted differences in means or odds ratios (ORs) with 95% CIs and p-values.

Repeated measures analyses will be conducted of all outcomes incorporating all values over the 12-month follow-up period (including those at months 3 and 9). Models will include an interaction term between treatment group and time to assess whether treatment effects are sustained or emerge later.

We will perform regression model diagnostics using graphs and alternative methods of analysis will be considered if the assumptions of the model are not met.

6.5 Subgroup analyses

Three pre-defined subgroup analyses will be carried out to assess the difference in treatment effect on the primary outcome according to the following characteristics assessed at baseline: chronicity and severity of depression and personality difficulties. In each case separately, effect modification will be assessed by including an interaction term in the regression model and formal tests of interaction will be performed to test whether the treatment effect differs between these groups. As the study was not powered to detect such effects, results will be interpreted with caution.

6.6 Exploratory analyses

If and when possible, we will undertake exploratory analyses using regression methods to examine the impact of process measures such as extent of involvement in CBT/other interventions (including use of INTERACT platform), use of online materials, and therapeutic alliance.

7. ANALYSIS OF SAFETY

As outlined in section 3.1, all adverse events and adverse reactions as described below will firstly be tabulated by allocated group then secondly by treatment received. Data will be collected from the point of consent (baseline assessment) until the 12-month follow-up assessment or point of withdrawal from the study. The number of events, number of patients having at least one event and the number of patients with more than one event will be tabulated. Serious adverse events will also be listed.

7.1 Adverse events

Adverse events (AEs) are any untoward medical occurrence in a study participant to whom an intervention has been administered. AEs that might be expected to occur at a higher rate in this group of participants include episodes of self-harm not requiring hospital admission and worsening of depression sufficient to require referral to a clinician. Although these AEs are expected, they will still be reported.

Variations in mood, including worsening of depression that does not lead to self-harm or hospitalisation, are commonly seen during therapy and would not be reported as individual adverse events.

There may be more AE reports in the intervention group as a result of their regular contact with the study team (i.e. receipt of up to 12 therapy sessions with the study therapist, and more regular completion of the PHQ-9 and related patient safety risk assessments as part of therapy). This will be taken into consideration when interpreting the pattern of AEs.

7.2 Serious adverse events

Serious adverse events (SAEs) are events that:

- result in death
- are life-threatening
- require hospitalisation or prolongation of existing hospitalisation

- result in persistent or significant disability or incapacity
- are otherwise considered medically significant by the investigator.

SAEs are "expected" if there is a more common occurrence in this study population regardless of the study itself. Expected SAEs in this study are listed below and will be reported:

- self-harm leading to hospitalisation
- suicidal attempts leading to hospitalisation
- worsening depression leading to hospitalisation

Admission to hospital for pre-planned surgery for pre-existing conditions will not be reported as an SAE.

7.3 Suspected Serious Adverse Reaction (SSAR)

These are any serious adverse event that is suspected (possibly/probably/definitely) to be related to the intervention.

7.4 Non-IMP Suspected Unexpected Serious Adverse Reaction (non-imp SUSAR)

An SAE that occurs in a non-IMP trial and is:

- "Related" – that has, possibly, probably or definitely resulted from administration of any of the research procedures, **and**
- "Unexpected" – that is, the type of event is not listed in the protocol (or above) as an expected occurrence.

All AEs will be assessed for seriousness, causality and expectedness by Prof David Kessler or nominated deputy clinician.

8. CHANGES TO THE SAP

All changes made to the planned statistical analyses are described below:

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

9. FINAL REPORT TABLES AND FIGURES (SUBJECT TO CHANGE)

Section	Outputs	
9.1 Population	**Tables, figures and listings detailing the study population**	
	Figure F1	Predicted and actual recruitment
	Table T1	Practice details by centre
	Figure F2	Flow of participants: recruitment pathway
	Figure F3	Flow of participants: randomisation onwards
	Table T2	Recruitment statistics by centre
	Table T3	Comparison of age and gender of those identified by GPs as potential participants and those who were excluded based on the record search
	Table T4	Comparison of age and gender of those accepting the postal invitation to participate in the trial with those who declined and those who did not respond
	Table T5	Reasons for declining to participate (postal invitation)
	Table T6	Comparison of age and gender of those undergoing telephone screening compared to those who declined or were unable to be contacted
	Table T7	Comparison of age and gender of those completing the baseline assessment and those declining to attend or not responding
	Table T8	Comparison of demographic characteristics and socio-economic status of those who were eligible and not eligible at baseline
	Table T9	Protocol deviations
	Table T10	Details of individual protocol deviations
	Table T11	Withdrawals from the trial
	Table T12	Details of individual withdrawals from the trial
9.2 Baseline data	**Summary tables of demographic information**	
	Table T13	Baseline comparability of randomised groups
	Table T14	Antidepressant medication use at baseline
	Table T15	Summary of baseline variables related to missing BDI-II at 6 months
	Table T16	Summary of baseline variables related to missing BDI-II at 12 months
9.3 Intervention delivery	**Summary tables of how the intervention was delivered**	
	Table T17	Therapist caseload
	Table T18	Number of CBT sessions attended
	Table T19	Completion rates of therapy across all centres
9.3 Outcomes	**Summary data and treatment estimates**	
	Table T20	Timing of questionnaire completion at all follow-ups
	Table T21	Primary outcome: mean and difference in mean BDI-II scores at 6 months
	Table T22	Means and differences in mean BDI-II scores at 12 months and repeated measures analysis
	Table T23	Percentage reduction in depressive symptoms on the BDI-II at 6 and 12 months
	Table T24	Percentage and OR of “response to treatment” (improvement of at least 50% in BDI-II score compared with baseline) at 6 and 12 months
	Table T25	Percentage and OR of “remission of symptoms” (BDI-II of less than 10) at 6 and 12 months
	Table T26	Means and differences in mean GAD-7 scores at 6 and 12 months

Section	Outputs	
	Table T27	Means and differences in mean PHQ-9 scores at 6 and 12 months
	Table T28	Means and differences in mean WSAS scores at 6 and 12 months
	Table T29	Comparison of results from ITT and CACE analyses for the primary outcome of BDI-II score at 6 months
	Table T30	Comparison of results from ITT and CACE analyses for the primary outcome of BDI-II score at 12 months
	Table T31	Comparison of results of ITT analysis of complete cases with ITT analysis where missing data were imputed using "better" and "worse" case scenarios and multiple imputation for primary outcome of BDI-II score at 6 months
	Table T32	Other therapies received during follow-up
9.4 Safety data	**Summary tables and listings of all adverse events and serious adverse events**	
	Table T33	Expected adverse events and serious adverse events
	Table T34	Expected adverse events and serious adverse events by relatedness of treatment (UR: un-related; RL: related)
	Table T35	Unexpected adverse events and serious adverse events
	Table T36	Details of serious unexpected adverse events
	Table T37	Number of adverse events per patient stratified by relatedness to treatment

9.1 Population

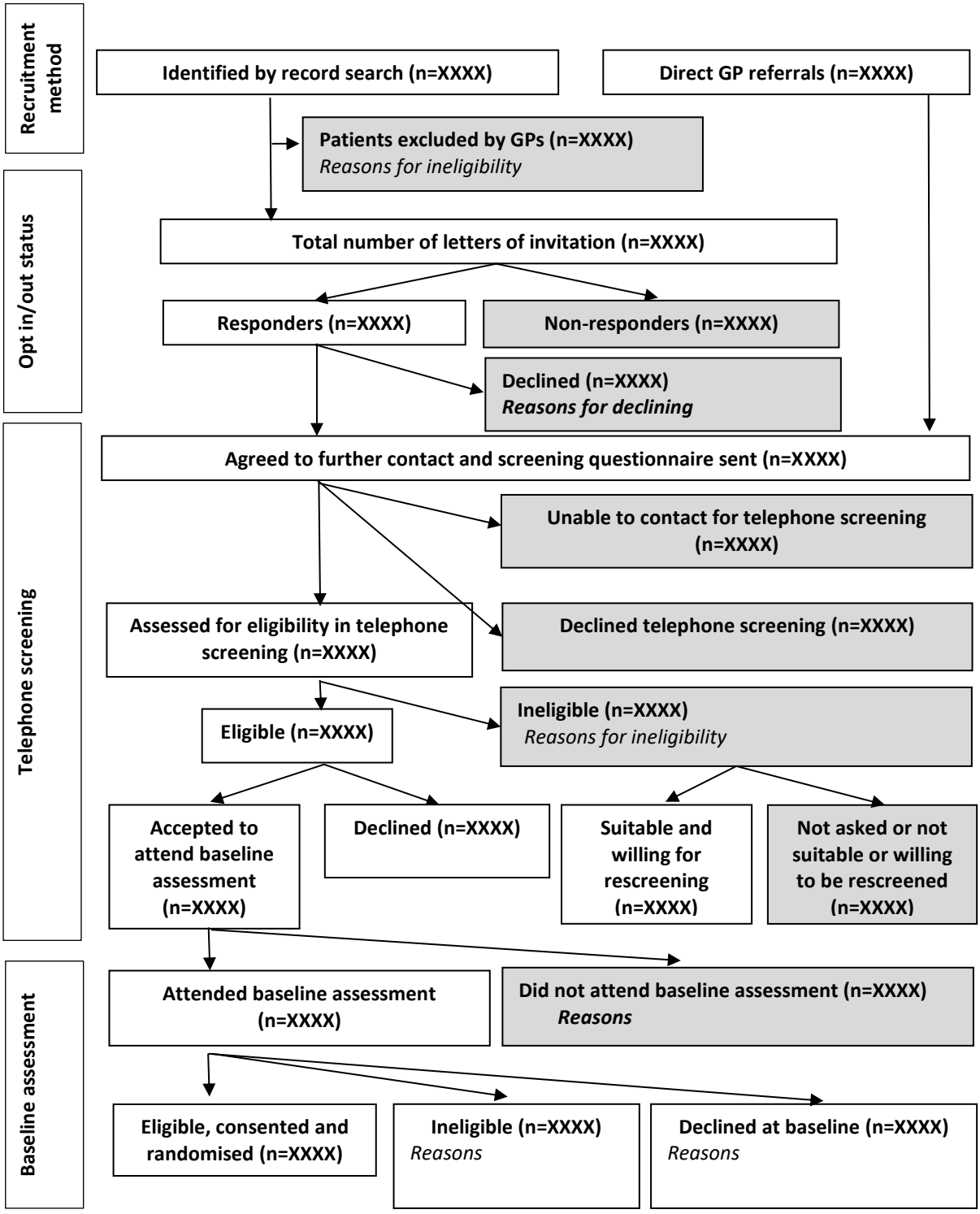
Figure F1 Predicted and actual recruitment

X axis: Month; Y axis: Number of patients recruited

Table T1 Practice details by centre

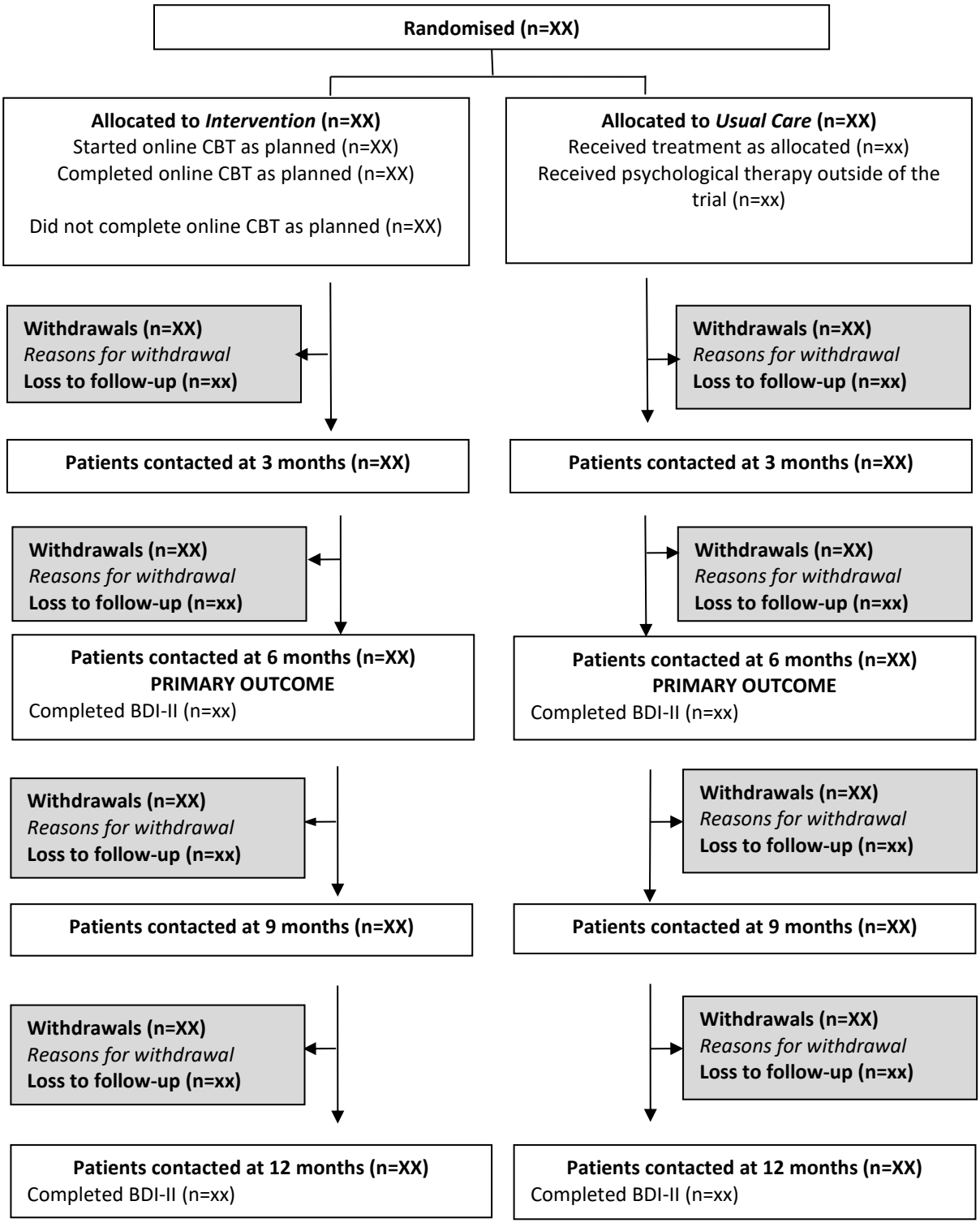
	Bristol	London	York	Total
Number of practices				
Practice list size: median (IQR)				
Number of full-time GPs per practice: mean (SD)				
Number of patients per practice				
Number of referrals: median (IQR)				
Number of invites mailed: median (IQR)				
Number of telephone screenings: median (IQR)				
Number of baseline assessments: median (IQR)				
Number of participants recruited: median (IQR)				

Figure F2 **Flow of participants: recruitment pathway**



Notes:
Some patients may be ineligible for more than one reason

Figure F3 **Flow of participants: randomisation onwards**



Notes:
Some patients may be ineligible for more than one reason

Table T2 **Recruitment statistics by centre**

	Bristol	**London**	**York**	**Total**
Number of practices				
Invitations and GP referrals				
Number of GP referrals				

Number of postal invitations sent				
Number of postal opt-ins received				
Number invited for telephone screening (percentage of those invited by post or referred)				
Telephone screening				
Number eligible for baseline assessment (percentage of those undergoing telephone screening)				
Baseline assessments				
Number attending baseline assessment (percentage of those eligible at telephone screening)				
Randomisations				
Number randomised (percentage of those attending baseline assessments)				

Table T3 Comparison of age and gender of those identified by GPs as potential participants and those who were excluded based on the record search

		Age			Female		
	N	**n[a]**	**Mean**	**SD**	**n[a]**	**N**	**%**
Excluded							
Potential participant							

[a] *Number with available data*

Table T4 Comparison of age and gender of those accepting the postal invitation to participate in the trial with those who declined and those who did not respond

		Age			Female		
	N	**n[a]**	**Mean**	**SD**	**n[a]**	**N**	**%**
Did not respond							
Declined							
Accepted							

[a] *Number with available data*

Table T5 Reasons for declining to participate (postal invitation)

Reason for declining	**Number of decliners**	**%**
….		

Table T6 Comparison of age and gender of those undergoing telephone screening compared to those who declined or were unable to be contacted

		Age			Female		
	N	**n[a]**	**Mean**	**SD**	**n[a]**	**N**	**%**
Did not complete telephone screening or were							

unable to be contacted							
Completed telephone screening assessment							

[a] *Number with available data*

Table T7 Comparison of age and gender of those completing the baseline assessment and those declining to attend or not responding

		Age			Female		
	N	**n[a]**	**Mean**	**SD**	**n[a]**	**N**	**%**
No (declined or not responding)							
Yes (agreed)							

[a] *Number with available data*

Table T8 Comparison of demographic characteristics and socio-economic status of those who were eligible and not eligible at baseline

	Ineligible at baseline	**Eligible (including declined to participate)**
Age; **n[a]** **mean (SD)**		
Female; **n[a]** **n (%)**		
Employment status **N** **n[a]**		
In paid employment (full/part-time); n (%)		
Not in employment; n (%)		
Unemployed owing to ill health; n (%)		
Educational attainment **N** **n[a]**		
A-level, higher grade or above; n (%)		
GCSE, standard grade or above; n (%)		
No formal qualifications; n (%)		
Housing **N** **n[a]**		
Home owner; n (%)		
Tenant or living with relative/friend; n (%)		
Hostel/care home, homeless or other; n (%)		

[a] *Number with available data*

Table T9 Protocol deviations

	Randomised to Intervention (n=)		Randomised to Usual care (n=)		Overall (n=)	
	Patients	**%**	**Patients**	**%**	**Patients**	**%**
Any protocol deviation						
…..						

Table T10 Details of individual protocol deviations

Allocated treatment group	**Centre**	**Further details (exact nature dependent upon type of deviation)**
…		

Table T11 Withdrawal from the trial

	Randomised to Intervention (n=XX)		Randomised to Usual care (n=XX)		Overall (n=XX)	
	n	**%**	**n**	**%**	**N**	**%**
Any withdrawal from the trial						
Reason						

Table T12 Details of individual withdrawals from the trial

Allocated treatment group	**Days between randomisation and withdrawal from the trial (estimated where dates not provided)**	**Patient withdrew consent or clinician's decision**
….		

9.2 Baseline data

Table T13 Baseline comparability of randomised groups

	Intervention (n=xx)	**Usual care (n=xx)**	**Total (n=xx)**
Stratification variable: centre n(%)			
Bristol London York			
Minimisation variables			
Female: n (%)			
Baseline BDI: n(%) ≤25 26-35 ≥36			
Currently using anti-depressants: n(%)			
Socio-demographic variables			
Age (years): mean (SD)			
Self-reported gender: n(%) Male Female			

Other			
Ethnic group, white: n(%)			
Marital status: n(%) Married/living as married Single Separated/widowed/divorced			
Employment status: n(%) In paid employment (full/part-time) Not in employment Unemployment due to ill health			
Educational attainment: n(%) A-level, higher grade or above GCSE, standard grade or above No formal qualifications			
Housing: n(%) Home owner Tenant or living with relative/friend Hostel/care home, homeless or other			
Financial well-being: n(%) Living comfortably/doing all right Just about getting by Finding it difficult/very difficult to make ends meet			
IT literacy: I am confident using computers; n(%) Strongly agree Agree Neither agree nor disagree Disagree Strongly disagree			
Alcohol consumption AUDIT-PC score: median (IQR)			
Number of life events in the past 6 months: mean (SD)			
Social support score: mean (SD)			
Long-standing illness or disability; n(%) Any Alzheimer's disease/dementia Angia/long-term heart problem Arthritis/long-term joint problem Asthma/long-term chest problem Blindness/severe visual impairment Cancer in the last 5 years Deafness/severe hearing impairment Diabetes Epilepsy High blood pressure Kidney or liver disease Learning difficulty Long-term back problem Long-term mental health problem Long-term neurological problem Another long-term condition None of the above			

Prefer not to say			
Work and social adjustment score (WSAS); mean (SD)			
Standardised assessment of personality scale (SAPS); mean (SD)			
Post-traumatic stress Ever experienced a traumatic event (serious accident/fire, physical/sexual assault/abuse, earthquake/flood, war, seeing someone be killed/seriously injured or having a loved one die through homicide/suicide); n(%) **Score; n(%)** 0 1 2 3 4 5			
Neurotic symptom score; mean (SD)			
Measures of depression			
Suffered depression in the past: n (%)			
Family history of depression: n (%)			
Previous referral to a psychiatrist for depression: n (%)			
Number of prior episodes of depression: n (%) 0-1 2-4 ≥5			
Among those currently taking anti-depressants – Length of current course of anti-depressants: n(%) Less than 6 weeks 6 weeks – 3 months 3-6 months 6-12 months More than 12 months			
Among those currently taking anti-depressants – Adherence to anti-depressants at baseline: n(%) I have taken my tablets every day I have taken my tablets nearly every day I have taken more than half my tablets I have taken less than half my tablets I have hardly taken any of my tablets I have not taken any of my tablets			
ICD-10 primary diagnosis: n(%) Mild Moderate Severe			
Secondary psychiatric diagnosis according to the CIS-R: n(%) *List according to responses*			

BDI-II score: mean (SD)			
GAD-7 score: mean (SD)			
PHQ-9 score: mean (SD)			
EQ-5D-5L score: mean (SD)			
CIS-R score: mean (SD)			
Suicidal ideation (CIS-R thoughts/plans): n (%)			

Note: Where data are incomplete for some variables, the numbers with information available are listed here

Table T14 Antidepressant medication use at baseline

	Intervention (n=xx)		**Usual care (n=xx)**		**Total (n=XXX)**	
Antidepressant medication	**n**	**%**	**n**	**%**	**n**	**%**
List according to reported data						

Table T15 Summary of baseline variables related to missing BDI-II data at 6 months

	Present (n=xx)	**Absent (n=xx)**	**Total (n=xx)**
Stratification variable: centre n(%)			
Bristol London York			
Minimisation variables			
Female: n (%)			
Baseline BDI: n(%) ≤25 26-35 ≥36			
Currently using anti-depressants: n(%)			
Socio-demographic variables			
Age (years): mean (SD)			
Self-reported gender: n(%) Male Female Other			
Ethnic group, white: n(%)			
Marital status: n(%) Married/living as married Single Separated/widowed/divorced			
Employment status: n(%) In paid employment (full/part-time) Not in employment Unemployment due to ill health			
Educational attainment: n(%) A-level, higher grade or above GCSE, standard grade or above No formal qualifications			
Housing: n(%) Home owner Tenant or living with relative/friend Hostel/care home, homeless or other			
Financial well-being: n(%) Living comfortably/doing all right Just about getting by			

Finding it difficult/very difficult to make ends meet			
IT literacy: I am confident using computers; n(%) Strongly agree Agree Neither agree nor disagree Disagree Strongly disagree			
Alcohol consumption AUDIT-PC score: median (IQR)			
Number of life events in the past 6 months: mean (SD)			
Social support score: mean (SD)			
Long-standing illness or disability; n(%) Any Alzheimer's disease/dementia Angia/long-term heart problem Arthritis/long-term joint problem Asthma/long-term chest problem Blindness/severe visual impairment Cancer in the last 5 years Deafness/severe hearing impairment Diabetes Epilepsy High blood pressure Kidney or liver disease Learning difficulty Long-term back problem Long-term mental health problem Long-term neurological problem Another long-term condition None of the above Prefer not to say			
Work and social adjustment score (WSAS); mean (SD)			
Standardised assessment of personality scale (SAPS); mean (SD)			
Post-traumatic stress Ever experienced a traumatic event (serious accident/fire, physical/sexual assault/abuse, earthquake/flood, war, seeing someone be killed/seriously injured or having a loved one die through homicide/suicide); n(%) **Score; n(%)** 0 1 2 3 4 5			
Neurotic symptom score; mean (SD)			

Measures of depression			
Suffered depression in the past: n (%)			
Family history of depression: n (%)			
Previous referral to a psychiatrist for depression: n (%)			
Number of prior episodes of depression: n (%) 0-1 2-4 ≥5			
Among those currently taking anti-depressants – Length of current course of anti-depressants: n(%) Less than 6 weeks 6 weeks – 3 months 3-6 months 6-12 months More than 12 months			
Among those currently taking anti-depressants – Adherence to anti-depressants at baseline: n(%) I have taken my tablets every day I have taken my tablets nearly every day I have taken more than half my tablets I have taken less than half my tablets I have hardly taken any of my tablets I have not taken any of my tablets			
ICD-10 primary diagnosis: n(%) Mild Moderate Severe			
Secondary psychiatric diagnosis according to the CIS-R: n(%) *List according to responses*			
BDI-II score: mean (SD)			
GAD-7 score: mean (SD)			
PHQ-9 score: mean (SD)			
EQ-5D-5L score: mean (SD)			
CIS-R score: mean (SD)			
Suicidal ideation (CIS-R thoughts/plans): n (%)			

Note: Where data are incomplete for some variables, the numbers with information available are listed here

Table T16 Summary of baseline variables related to missing BDI-II data at 12 months

	Present (n=xx)	**Absent (n=xx)**	**Total (n=xx)**
Stratification variable: centre n(%)			
Bristol London York			
Minimisation variables			
Female: n (%)			
Baseline BDI: n(%) ≤25			

26-35 ≥36			
Currently using anti-depressants: n(%)			
Socio-demographic variables			
Age (years): mean (SD)			
Self-reported gender: n(%) Male Female Other			
Ethnic group, white: n(%)			
Marital status: n(%) Married/living as married Single Separated/widowed/divorced			
Employment status: n(%) In paid employment (full/part-time) Not in employment Unemployment due to ill health			
Educational attainment: n(%) A-level, higher grade or above GCSE, standard grade or above No formal qualifications			
Housing: n(%) Home owner Tenant or living with relative/friend Hostel/care home, homeless or other			
Financial well-being: n(%) Living comfortably/doing all right Just about getting by Finding it difficult/very difficult to make ends meet			
IT literacy: I am confident using computers; n(%) Strongly agree Agree Neither agree nor disagree Disagree Strongly disagree			
Alcohol consumption AUDIT-PC score: median (IQR)			
Number of life events in the past 6 months: mean (SD)			
Social support score: mean (SD)			
Long-standing illness or disability; n(%) Any Alzheimer's disease/dementia Angia/long-term heart problem Arthritis/long-term joint problem Asthma/long-term chest problem Blindness/severe visual impairment Cancer in the last 5 years Deafness/severe hearing impairment Diabetes Epilepsy			

High blood pressure Kidney or liver disease Learning difficulty Long-term back problem Long-term mental health problem Long-term neurological problem Another long-term condition None of the above Prefer not to say			
Work and social adjustment score; mean (SD)			
Standardised assessment of personality scale; mean (SD)			
Post-traumatic stress Ever experienced a traumatic event (serious accident/fire, physical/sexual assault/abuse, earthquake/flood, war, seeing someone be killed/seriously injured or having a loved one die through homicide/suicide); n(%) **Score; n(%)** 0 1 2 3 4 5			
Neurotic symptom score; mean (SD)			
Measures of depression			
Suffered depression in the past: n (%)			
Family history of depression: n (%)			
Previous referral to a psychiatrist for depression: n (%)			
Number of prior episodes of depression: n (%) 0-1 2-4 ≥5			
Among those currently taking anti-depressants – Length of current course of anti-depressants: n(%) Less than 6 weeks 6 weeks – 3 months 3-6 months 6-12 months More than 12 months			
Among those currently taking anti-depressants – Adherence to anti-depressants at baseline: n(%) I have taken my tablets every day I have taken my tablets nearly every day I have taken more than half my tablets I have taken less than half my tablets			

I have hardly taken any of my tablets I have not taken any of my tablets			
ICD-10 primary diagnosis: n(%) Mild Moderate Severe			
Secondary psychiatric diagnosis according to the CIS-R: n(%) *List according to responses*			
BDI-II score: mean (SD)			
GAD-7 score: mean (SD)			
PHQ-9 score: mean (SD)			
EQ-5D-5L score: mean (SD)			
CIS-R score: mean (SD)			
Suicidal ideation (CIS-R thoughts/plans): n (%)			

Note: Where data are incomplete for some variables, the numbers with information available are listed here

9.3 Intervention delivery

Table T17 Therapist caseload

Therapist ID	N	%

Table T18 Number of CBT sessions attended

Number of sessions attended	n	%
..		

Table T19 Completion rates for therapy across all centres

Therapy outcome	Bristol	London	York	Total
Discharged for non-adherence to the intervention				
Withdrew from therapy				
Agreed end				

9.4 Outcomes

Table T20 Timing of questionnaire completion at all follow-ups

Follow-up	n	Mean (months)	SD	Number within 2 months	Proportion within 2 months
6 months					
12 months					

Table T21 Primary outcome: mean and difference in mean BDI-II scores at 6 months

Randomisation groups	n	Mean	SD	Difference in means[a]	95% CI	p-value	Difference in means[b]	95% CI	p-value
Intervention									
Usual care									

Total N									

[a] *ITT analysis adjusted for baseline BDI-II score and the stratification and other minimisation variables*
[b] *ITT analysis additionally adjusted for additional variables that show an imbalance between treatment groups at baseline*

Table T22 Means and differences in mean BDI-II scores at 12 months and repeated measures analysis

Randomisation groups	n	Mean	SD	Difference in means[a]	95% CI	p-value	Difference in means[b]	95% CI	p-value
Intervention									
Usual care									
Total N									
Repeated measures analysis									

[a] *ITT analysis adjusted for baseline BDI-II score and the stratification and other minimisation variables*
[b] *ITT analysis additionally adjusted for additional variables that show an imbalance between treatment groups at baseline*

Table T23 Percentage reduction in depressive symptoms on the BDI-II at 6 and 12 months

	N	Difference in percent reduction[a]	95% CI	p-value	Difference in percent reduction[b]	95% CI	p-value
6 months follow-up							
12 months follow-up							

[a] *ITT analysis adjusted for baseline BDI-II score and the stratification and other minimisation variables*
[b] *ITT analysis additionally adjusted for additional variables that show an imbalance between treatment groups at baseline*

Table T24 Percentage and OR of "response to treatment" (improvement of at least 50% in BDI-II score compared with baseline) at 6 and 12 months

	Follow-up					
	6 months (n=)			12 months (n=)		
	N	n	%[a]	N	n	%[a]
Intervention						
Usual care						
Regression analyses						
	N	OR[b]	95% CI	p-value	OR[c]	95% CI	p-value
6 months follow-up							
12 months follow-up							
Repeated measures							

[a] *Number of patients reporting an improvement of at least 50% in BDI-II score compared with baseline (n) as a percentage of the total number (N) in the group*
[b] *ITT analysis adjusted for baseline BDI-II score and the stratification and other minimisation variables*
[c] *ITT analysis additionally adjusted for additional variables that show an imbalance between treatment groups at baseline*

Table T25 Percentage and OR of "remission of symptoms" (BDI-II of less than 10) at 6 and 12 months

	Follow-up					
	6 months (n=)			12 months (n=)		
	N	n	%[a]	N	n	%[a]
Intervention						
Usual care						
Regression analyses						
	N	OR[b]	95% CI	p-value	OR[c]	95% CI	p-value
6 months follow-up							
12 months follow-up							
Repeated measures							

[a] *Number of patients reporting a BDI-II of less than 10 (n) as a percentage of the total number (N) in the group*
[b] *ITT analysis adjusted for baseline BDI-II score and the stratification and other minimisation variables*
[c] *ITT analysis additionally adjusted for additional variables that show an imbalance between treatment groups at baseline*

Table T26 Means and differences in mean GAD-7 scores at 6 and 12 months

	Follow-up											
	3 months (n=)			6 months (n=)			9 months (n=)			12 months (n=)		
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Intervention												
Usual care												
Regression analyses												
	N	Difference in means[a]	95% CI	p-value	Difference in means[b]	95% CI	p-value					
6 months follow-up												
12 months follow-up												
Repeated measures including values at 3, 6, 9 and 12 months												

[a] *ITT analysis adjusted for baseline GAD-7 score and the stratification and other minimisation variables*
[b] *ITT analysis additionally adjusted for additional variables that show an imbalance between treatment groups at baseline*

Table T27 Means and differences in mean PHQ-9 scores at 6 and 12 months

	Follow-up											
	3 months (n=)			6 months (n=)			9 months (n=)			12 months (n=)		
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Intervention												
Usual care												
Regression analyses												

	N	Difference in means[a]	95% CI	p-value	Difference in means[b]	95% CI	p-value
6 months follow-up							
12 months follow-up							
Repeated measures including values at 3, 6, 9 and 12 months							

[a] *ITT analysis adjusted for baseline PHQ-9 score and the stratification and other minimisation variables*
[b] *ITT analysis additionally adjusted for additional variables that show an imbalance between treatment groups at baseline*

Table T28 Means and differences in mean WSAS scores at 6 and 12 months

	Follow-up					
	6 months (n=)			12 months (n=)		
	n	Mean	SD	N	Mean	SD
Intervention						
Usual care						
Regression analyses						

	N	Difference in means[a]	95% CI	p-value	Difference in means[b]	95% CI	p-value
6 months follow-up							
12 months follow-up							
Repeated measures							

[a] *ITT analysis adjusted for baseline WSAS score and the stratification and other minimisation variables*
[b] *ITT analysis additionally adjusted for additional variables that show an imbalance between treatment groups at baseline*

Table T29 Comparison of results from ITT and CACE analyses for the primary outcome of BDI-II score at 6 months

	n	Difference in means[a]	95% CI	p-value
ITT				
CACE – attending at least 6 sessions or reaching a jointly agreed end to therapy				

[a] *Adjusted for baseline BDI-II score, stratification and other minimisation variables and additional variables showing an imbalance between treatment groups at baseline*

Table T30 Comparison of results from ITT and CACE analyses for the outcome of BDI-II score at 12 months

	n	Difference in means[a]	95% CI	p-value
ITT				
CACE – attending at least 6 sessions or				

reaching a jointly agreed end to therapy				

[a] *Adjusted for baseline BDI-II score, stratification and other minimisation variables and additional variables showing an imbalance between treatment groups at baseline*

Table T31 Comparison of results of ITT analysis of complete cases with ITT analysis where missing data were imputed using "better" and "worse" case scenarios and multiple imputation for primary outcome of BDI-II score at 6 months

	n	Difference in means[a]	95% CI	p-value
Complete case				
Better case scenario				
Worse case scenario				
Multiple imputation				

[a] *Adjusted for baseline BDI-II score, stratification and other minimisation variables and additional variables showing an imbalance between treatment groups at baseline*

Table T32 Other therapies received during follow-up

Type of therapy received	Randomised to Intervention	Randomised to usual care
..		

9.5 Safety data

Table T33 Expected adverse events and serious adverse events

	Allocated to Intervention (n=XX patients)						Allocated to usual care (n=XX patients)					
	Events		Patients (n=XX)				Events		Patients (n=XX)			
	AE	SAE	AE	%	SAE	%	AE	SAE	AE	%	SAE	%
PATIENTS WITH ≥1 EVENT												
Details by type												
	Received intervention (n=xx patients)						Received usual care (n=XX patients)					
	Events		Patients (n=XX)				Events		Patients (n=XX)			
	AE	SAE	AE	%	SAE	%	AE	SAE	AE	%	SAE	%
PATIENTS WITH ≥1 EVENT												
Details by type												

Notes: *SAEs are a subset of AEs.*

Table T34 Expected adverse events and serious adverse events by relatedness to treatment (UR: un-related; RL: related)

	Received Intervention												Received usual care											
	Events				Patients (n=XX)								Events				Patients (n=XX)							
	AE		SAE		AE				SAE				AE		SAE		AE				SAE			
	UR	RL	UR	RL	UR	%	RL	%	UR	%	RL	%	UR	RL	UR	RL	UR	%	RL	%	UR	%	RL	%
PATIENTS WITH ≥1 EVENT																								
Details by type																								
	Received Intervention												Received usual care											
	Events				Patients (n=XX)								Events				Patients (n=XX)							
	AE		SAE		AE				SAE				AE		SAE		AE				SAE			
	UR	RL	UR	RL	UR	%	RL	%	UR	%	RL	%	UR	RL	UR	RL	UR	%	RL	%	UR	%	RL	%
PATIENTS WITH ≥1 EVENT																								

	Received Intervention												Received usual care											
	Events				Patients (n=XX)								Events				Patients (n=XX)							
	AE		SAE		AE				SAE				AE		SAE		AE				SAE			
	UR	RL	UR	RL	UR	%	RL	%	UR	%	RL	%	UR	RL	UR	RL	UR	%	RL	%	UR	%	RL	%
Details by type																								

Notes: *SAEs are a subset of AEs.*

Table T35 Unexpected adverse events and serious adverse events

		Allocated to Intervention (n=XX)		Allocated to usual care (n=XX)		Received Intervention (n=XX)		Received usual care (n=XX)	
		n	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Number of patients experiencing one or more SAEs									
Number of events									
Reason event classified as SAE	Resulted in death								
	Is/was life threatening								
	Resulted in persistent or significant disability/incapacity								
	Prolonged ongoing hospitalisation/ caused hospitalisation (other than hospitalisations for social reasons in absence of an adverse event, in-clinic protocol measures and surgery/procedure planned before entry into the trial)								
	Other								
Relatedness to *treatment*	Not related								
	Unlikely to be related								
	Possibly related								
	Probably related								
	Definitely related								

Table T36 Details of serious unexpected adverse events

Study ID=	Treatment randomised to=	Treatment received=	Patient withdrawn from study (and when)=
Treatment start date=	Timing of SAE in terms of starting therapy =		
Brief description of event=	Location=	Maximum intensity=	Relatedness=
SAE start date=	SAE resolution date=	Event resulted in death=	Event was life threatening=
Event resulted in persistent/significant disability/incapacity=	Event prolonged ongoing hospitalisation/resulted in hospitalisation=	Other reason for reporting as SAE (with details)=	

Table T37 Number of adverse events per patient stratified by relatedness to treatment

	Allocated to Intervention (n=XX)						**Allocated to usual care (n=XX)**					
	All		**Un-related**		**Related**		**All**		**Un-related**		**Related**	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0												
1												
2												
3												
4+												
Total												
	Received Intervention (n=XX)						**Received usual care (n=XX)**					
	All		**Un-related**		**Related**		**All**		**Un-related**		**Related**	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0												
1												
2												
3												
4+												
Total												

BIBLIOGRAPHY

1. Radhakrishnan M, Hammond G, Jones PB, Watson A, McMillan-Shields F, Lafortune L. Cost of improving Access to Psychological Therapies (IAPT) programme: an analysis of cost of session, treatment and recovery in selected Primary Care Trusts in the East of England region. Behav Res Ther. 2013;51(1):37-45.
2. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev. 2012;32(4):329-42.
3. Helgadottir FM, R.; Onslow, M.; Packman, A.; O'Brian, S. . Online CBT I: Bridging the Gap Between Eliza and Modern Online CBT Treatment Packages. Behaviour Change. 2009;26:245-53.
4. Beck AS, R.; Brown, G. Manual for the Beck Depression Inventory-II. San Antonio, Texas: Psychological Corporation; 1996.
5. Lewis G. Assessing psychiatric disorder with a human interviewer or a computer. J Epidemiol Community Health. 1994;48(2):207-10.
6. Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. Psychol Med. 1992;22(2):465-86.
7. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606-13.
8. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166(10):1092-7.
9. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
10. Marks I. Behavioural Psychotherapy. Bristol: John Wright; 1986 1986.
11. John OD, E.M.; Kentle, R.L. The Big Five Inventory - Versions 4a and 54. Berkeley, California: University of California, Berkeley, Institute of Personality and Social Research; 1991.
12. Moran P, Leese M, Lee T, Walters P, Thornicroft G, Mann A. Standardised Assessment of Personality - Abbreviated Scale (SAPAS): preliminary validation of a brief screen for personality disorder. Br J Psychiatry. 2003;183:228-32.
13. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. Addiction. 1993;88(6):791-804.
14. The Primary Care PTSD Screen for DSM-5 (PC-PTSD-5) [Measurement instrument].
15. Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered Internet psychotherapy for depression in primary care: a randomised controlled trial. Lancet. 2009;374(9690):628-34.
16. Wiles N, Thomas L, Abel A, Ridgway N, Turner N, Campbell J, et al. Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care based patients with treatment resistant depression: results of the CoBalT randomised controlled trial. Lancet. 2013;381(9864):375-84.
17. Roberts C, Roberts SA. Design and analysis of clinical trials with clustering effects due to treatment. Clin Trials. 2005;2(2):152-62.
18. Chalder M, Wiles NJ, Campbell J, Hollinghurst SP, Searle A, Haase AM, et al. A pragmatic randomised controlled trial to evaluate the cost-effectiveness of a physical activity intervention as a treatment for depression: the treating depression with physical activity (TREAD) trial. Health Technol Assess. 2012;16(10):1-164, iii-iv.
19. Phillips R, Hazell L, Sauzet O, Cornelius V. Analysis and reporting of adverse events in randomised controlled trials: a review. BMJ Open. 2019;9(2):e024537.
20. Position statement on use of the EQ-5D-5L value set for England (updated October 2019): National Institute for Health and Care Excellence; 2019 [Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l.