Remote Approaches to Psychosocial Intervention Delivery (RAPID) Trial: a multi-arm, multi-stage randomised controlled trial

Statistical Analysis Plan
Version 2.0:
Date of version: 20/1/26

Based on Protocol Version 9.0 (dated 7/6/24)

ISRCTN: 33079589

Authors:
Professor Richard Emsley – senior trial statistician
Dr Alex Spiers – trial statistician
Mr Callum Glen – trial statistician

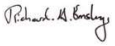
King's Clinical Trials Unit &
Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London

Signatories:

Junior Trial Statistician: Callum Glen (N/A as edits made after interim analysis)

~~Signed: Date:~~

Senior Trial Statistician: Professor Richard Emsley

Signed:  Date: 25/1/26

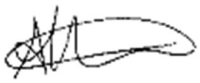
Independent DMC Statistician: Professor James Wason (N/A as edits made after interim analysis)

~~Signed: Date:~~

Independent TSC Statistician: Professor John Norrie

Signed:  Date: 2026-07-02

Chief Investigator: Professor Tony Morrison

Signed:  Date: 25/1/26

Approval and version control of the SAP

The initial signed off version of the SAP will be numbered as Version 1.0.
Minor amendments following this should be numbered as Version 1.1, Version 1.2, etc.
Major amendments should be numbered as Version 2.0, Version 3.0, etc.

After the initial SAP has been approved and signed, if any changes are required, these will be decided by the blinded Senior Trial Statistician so as to avoid any chance or appearance of introducing bias having seen unblinded data. A copy of the tracked changes version will be retained, and the updated version will be approved by the Chief Investigator, Senior Trial Statistician, and Independent TSC Statistician. Any changes between versions will be outlined below in Table 1.

Table 1. Amendments to the SAP

Version	Date of approval	Summary of changes
1.0	25/04/24	Original version
2.0	25/1/26	Updated to reflect interim analysis outcome and removal of Brighterside during Stage 1 Clarification about missing data from HER Added scoring rules for all scale measures

1 Description of the trial 6
1.1 Principal research objectives to be addressed 6
1.1.1 Primary and Secondary Objectives 6
1.2 Trial design including blinding 7
1.2.1 Figure 1 – RAPID Trial Flow Chart (showing original 4 arms) 9
1.3 Method of allocation of groups 9
1.4 Duration of the intervention period 10
1.5 Frequency and duration of follow-up 10
1.6 Visit windows 10
1.7 Data collection 10
1.7.1 Eligibility screening 10
1.7.2 Primary Outcome Measure 11
1.7.3 Secondary Outcome Measures 11
1.7.4 Baseline and demographic variables 12
1.7.5 Adverse events 12
1.8 Sample size estimation 13
1.9 Brief description of proposed analyses 14
2 Data analysis plan – Data description 14
2.1 Recruitment and representativeness of recruited patients 14
2.1.1 Figure 2 - Template CONSORT diagram for RAPID trial 16
2.2 Baseline comparability of randomised groups 17
2.3 Adherence to allocated treatment and treatment fidelity 17
2.4 Loss to follow-up and other missing data 17
2.5 Adverse event reporting 17
2.6 Assessment of outcome measures (unblinding) 18
2.7 Descriptive statistics for outcome measures 18
3 Data analysis plan – Stage 1 interim analysis 18
3.1 Decision rules 18
3.2 Statistical model 19
4 Data analysis plan – Stage 2 inferential analysis 19
4.1 Primary analysis description 19
4.1.1 Definition of primary estimand 19
4.1.2 Further details on handling of intercurrent events 20
4.1.3 Intention to treat population set 20
4.1.4 Analysis of primary outcome 20
4.1.5 Analysis of secondary outcomes 21
4.2 Statistical considerations 22
4.2.1 Rationale for using maximum likelihood estimation 22
4.2.2 Time points 22
4.2.3 Stratification and clustering 22
4.2.4 Missing items in scales and subscales 22
4.2.5 Missing baseline data 23
4.2.6 Missing outcome data 23
4.2.7 Method for handling multiple comparisons 23
4.2.8 Method for handling non-compliance (per protocol/CACE analyses) 23
4.2.9 Model assumption checks 24

4.2.10 Planned subgroup analyses 24
4.3 Software 24
5 Reference list 25
Appendix 1: Schedule of assessments and measures 26
Appendix 2: Scoring guidelines of clinical secondary outcome measures..... 27
Columbia Suicide Severity Rating Scale 27
Process Recovery Questionnaire (QPR) 27
Recovering Quality of Life (ReQoL-10) 27
Depression (PHQ-9) 28
Anxiety (GAD-7) 28
Adult HOPE Scale (AHS) 28
Entrapment Scale Short Form (4-item scale) 28
Adverse Effects – Learning From You 28
EQ-5D-5L 29
Appendix 3: Summary of blinding/unblinding..... 30

RAPID Statistical Analysis Plan

Chief Investigator: Professor Anthony P Morrison

Trial Coordinators: Dr Melissa Pyle
Dr Lucy Loftus

Senior statistician: Professor Richard Emsley

Junior statistician: Mr Callum Glen

NOTE: this version of the SAP is drafted based on Protocol version 9.0, which is for the 3-arm 2-stage design after the Brighterside arm has been dropped from the study and following the interim analysis resulted in PREVAIL being dropped.

1 Description of the trial

The RAPID trial aims to determine the clinical and cost effectiveness of two brief and remotely delivered psychosocial treatments, compared to treatment as usual, for people with serious mental health problems (SMHPs) who have had recent suicidal thoughts or behaviours. The psychosocial treatments being evaluated are:

- SAFETEL: a brief, innovative, evidence-based, telephone- or video-conferencing-delivered, clinical intervention that is implemented by assistant psychologists.
- PREVAIL: an intervention remotely delivered by Peer Support Workers (PSWs) by choice of telephone or video-conferencing consisting of up to 12 sessions over 12 weeks.

A previous intervention arm was dropped earlier in the study (April 2023) because of external evidence that the BrighterSide mobile app had a lack of efficacy on suicidal ideation. BrighterSide is an interactive self-help smartphone app for adults designed to help those with suicidal thinking to understand their thoughts and develop the best skills and strategies to help manage them.

1.1 Principal research objectives to be addressed

1.1.1 Primary and Secondary Objectives

Our primary objective is to answer the question of which brief, remote psychosocial intervention for people with SMHP who report recent suicidal ideation, or a suicide attempt is most clinically effective and cost-effective in preventing avoidable psychiatric hospital admissions in comparison to treatment as usual (TAU), and to determine the safety of the interventions.

Specific hypotheses are that, compared to TAU, our brief, remote interventions (SAFETEL, PREVAIL) plus TAU will lead to:

1) reduction in psychiatric hospital admissions over 6 months (primary outcome);
2) reduction in psychiatric hospital admissions over 3 months;
3) reduction in suicidal ideation over 3 and 6 months (key secondary outcome);
4) improvement in user-defined recovery and quality of life over 3 and 6 months.

We also hypothesise that our interventions will be cost-effective over 6 months in comparison to treatment as usual (TAU).

1.2 Trial design including blinding

RAPID was initially a 4-arm, 2-stage, multi-centre, parallel-group, randomised controlled trial recruiting people with SMHP diagnosis and current suicidal ideation. The design follows the methods outlined for MAMS designs with binary outcomes (Bratton et al). In Stage 1, the initial randomisation ratio was 1:1:1:2 in favor of TAU.

The design of the study was changed from a four-arm trial to a three-arm trial following removal of one arm, namely the BrighterSide mobile app intervention arm. This decision was based on, at the time, unpublished efficacy data from a large Australian RCT of the app which found no significant difference between those who received the app and the control group on the primary outcome of suicidal ideation. Based on the BrighterSide RCT results and early engagement data with the app from RAPID participants, the decision was made in conjunction with the RAPID trial independent DMC, independent chair and statistician from the TSC and the funder to remove the BrighterSide app arm.

As such the following changes were made to the design:

i. Randomisation ratio changed from 1:1:1:2 favouring TAU to 2:2:3 favouring TAU.
ii. Target sample size changed from 1235 to 1064 (not inclusive of the 54 participants randomised to BrighterSide at the time of dropping this arm).
iii. Number required for the interim analysis will change from the first 559 participants recruited with available primary outcome data to 385 participants with primary outcome data and provide 92% power.
iv. Removal of the following exclusion criterion since it was relevant only to the BrighterSide app, *"visual impairment, severe enough to prevent engagement with the BrighterSide app as provision would be impossible on both financial and logistical ground".*
v. The RAPID trial was initially classified as a clinical trial of a medical device by the MHRA based on the inclusion of the BrighterSide mobile app. Following removal of this intervention and on exit of the last participant allocated to BrighterSide in November 2023, the trial ceased to be a clinical trial of a medical device and under the regulation of the MHRA.

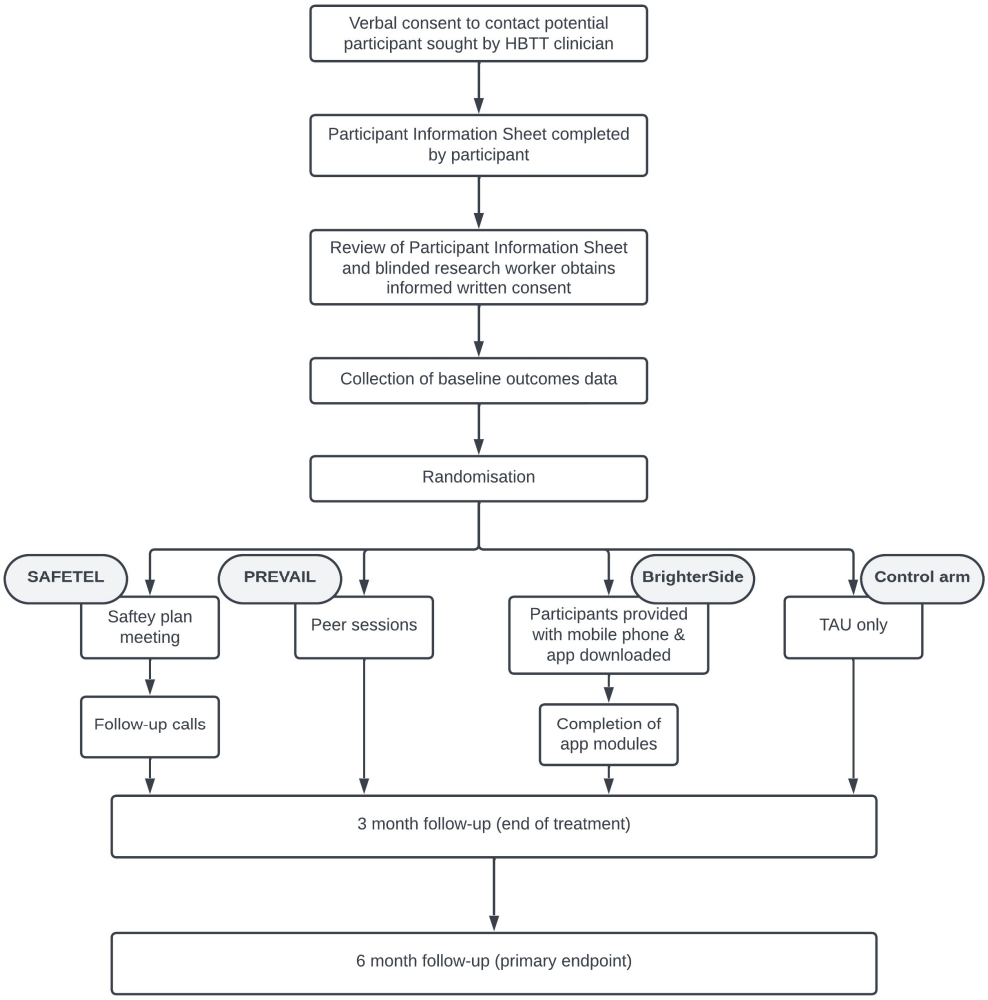
An interim analysis will be conducted at the end of Stage 1. Based on this analysis (decision rules are defined in **section 3.1**), the decision will made to:

- drop the **least** promising intervention arm and one intervention arm continues to Stage 2 **or**
- continue evaluating and recruiting into **both** intervention arms if both give promising results at the interim stage **or**
- drop both interventions and the trial stops recruiting.

Outcome measure data will be collected from participants at baseline, and subsequently at 3 months (end of treatment) and 6 months (follow-up). RAPID uses the same primary outcome at both stages of the trial and has a primary endpoint of 6 months after recruitment.

The trial statistician will remain blind to treatment allocation until the SAP has been finalised and approved; the senior trial statistician will remain blind until all analysis has been completed. Interim analysis of the primary outcome (admissions) will be carried out by the unblinded trial statistician and checked by an independent blinded trial statistician. Final analyses will be carried out by the trial statistician and overseen by the senior trial statistician.

1.2.1 Figure 1 – RAPID Trial Flow Chart (showing original 4 arms)



1.3 Method of allocation of groups

In Stage 1, randomisation will be in the ratio 2:2:3 in favour of TAU to the intervention groups and will be stratified by centre. Randomisation (at the individual level) will be independent and concealed, using randomised-permuted blocks of random size, which will be administered via a study-specific web-based portal developed by KCL Clinical Trials Unit (KCTU).

At the end of Stage 1, an interim analysis will be conducted to assess which interventions are the most promising based on available evidence collected thus far.
Based on the criteria outlined in **section 3.1**, in Stage 2 we will either:

- Drop the least promising intervention arm, with subsequent change to the allocation ratio of future participants to 1:1
- In the case of both intervention arms give promising results at interim stage, we will continue to allocate participants in the ratio 2:2:3 in favour of TAU

- drop both interventions and the trial stops recruiting.

We will prepare two randomisation systems so we can switch seamlessly to the new system without pausing recruitment following the interim analysis.

1.4 Duration of the intervention period

SAFETEL is a telephone- or video-conferencing software-delivered intervention that consists of an initial co-produced safety plan and follow-up contacts taking place over a 12-week period.

PREVAIL consists of 12 remotely delivered sessions over 12 weeks.

1.5 Frequency and duration of follow-up

Participants will complete measures at baseline and then subsequently follow-up measures at 3 months (end of treatment) and 6 months (primary endpoint) post-randomisation.

1.6 Visit windows

Follow-up assessments will occur at 3 months, after end of treatment and after 6 months.

1.7 Data collection

Treatment effectiveness will be assessed in terms of reduction in hospital admission. Therapeutic improvement will also be assessed using rate and degree of suicidal ideation, recovery from psychosis, hope and overall health status and utility.

Psychiatric hospital admission over 6 months (primary outcome) will be taken from participant electronic medical records.

All other outcome measures will be administered at baseline and subsequently at 3 months (end of treatment) and 6 months (follow-up) by RAs who will have been trained in the use of all the instruments and scales, to achieve a satisfactory level of inter-rater reliability.

Regular training sessions including the use of video and role-play will be conducted with all RAs in order to maintain reliability and prevent rater drift. Participants will be offered choices regarding length of assessments, including the option of breaks and multiple occasions. Assessment measures will be clearly prioritised so that the most important will be collected first to minimise missing data. Assessors will be blind to treatment condition.

1.7.1 Eligibility screening

1.7.1.1 Inclusion criteria

Our inclusion criteria are as follows:

1. Currently receiving care from a Home-Based Treatment Team/ Crisis team or have

done so within the last 14 days, since referrals to HBTT/Crisis team are associated with increased risk of a psychiatric hospital admission in the near future.
2. Aged 16+
3. Meet criteria for a diagnosis of SMHP (schizophrenia spectrum, bipolar, major depressive disorder, Emotionally Unstable Personality Disorder (EUPD), PTSD or cPTSD) since these diagnoses account for the majority of PHAs for mental health difficulties.
4. Experienced suicidal ideation or attempt within the last month / current crisis episode, as operationalised by answering 'yes' to items 1 or 2 of the Columbia-Suicide Severity Rating Scale.
5. Able to provide informed consent.
6. Receiving care from a Community Mental Health Team or Early Intervention Service, to ensure ongoing specialist mental health support following discharge from HBTT.

1.7.1.2 Exclusion criteria

1. Organic impairment, as this could be the cause of mental health symptoms rather than a SMHP.
2. Non-English speaking, since two of the interventions are remotely delivered talking therapies and one of the interventions is a smartphone app which has only been developed in English. Provision for non-English speakers would be impossible on both financial and logistical grounds.
3. Primary diagnosis of a drug or alcohol dependence, as this could be the cause of mental health symptoms rather than a SMHP.
4. Moderate to severe learning disability as confirmed by the participant's responsible clinician in their care team.
5. For both ethical and safety reasons, immediate risk to others as confirmed by the participant's responsible clinician in their care team.
6. Currently receiving psychiatric inpatient care (since people in recent contact with crisis teams may have already been admitted to hospital).

1.7.2 Primary Outcome Measure

The primary outcome is any admission to a psychiatric hospital admission over six months (treated as a dichotomous yes/no variable) which will be ascertained by screening the participants electronic patient records (EPR) for the six-month period.

1.7.3 Secondary Outcome Measures

(for references see trial protocol)

- Key secondary: Suicidal ideation as measured by the Columbia–Suicide Severity Rating Scale or C-SSRS. Derived outcomes used from the scale will be suicidal ideation score (maximum category suicidal ideation) and suicidal ideation intensity.
- Suicide attempt in last six months (baseline and T2 only) as measured by C-SSRS (ibid).
- The Process of Recovery Questionnaire, a 15-item service-user defined measure of recovery.

- The Recovering Quality of Life (ReQoL-10) measure focuses on aspects of recovery and quality of life and was designed for use in a broad range of mental health conditions. It was collaboratively developed with service users and clinicians.
- Anxiety will be measured using the GAD-7.
- Depression will be measured using the PHQ-9.
- Adult Hope Scale (AHS), an 8-item self-report questionnaire. We will only report total AHS score.
- Entrapment will be measured using the 4-item self-report Entrapment Scale Short-Form.
- Adverse effects measures on exit from the study at 6-month assessment using the Learning From You instrument.
- EQ-5D-5L, a generic and validated questionnaire for health-related quality of life, used to calculate quality-adjusted life years (QALYs) suitable for economic evaluation **(not analysed by trial statistician)**.
- Wider NHS and social care use will be assessed by an economic patient questionnaire which will be adapted from questionnaires used in similar populations **(not analysed by trial statistician)**.

Information on how secondary measures will be scored, along with how item-level missing data is treated, is described in Appendix E.

1.7.4 Baseline and demographic variables

- Date of birth
- Gender identity
- Highest level of education
- Employment status
- Marital status
- Living arrangements
- Ethnicity
- Religion / belief
- Serious mental health diagnosis (SMHP):
    - Schizophrenia
    - Bipolar disorder
    - Major depressive disorder
    - Emotionally unstable personality disorder

1.7.5 Adverse events

Adverse events (AEs), serious adverse events (SAEs) will be collected for all trial subjects from the time of their enrolment into the study. The time of enrolment is defined as the time at which, following recruitment, a subject sign and dates the informed consent form. Research workers from each site will note AEs/SAEs at each follow up interview and enter

these into the MACRO database. Any SAEs or suspected SAEs that are recorded will be reported to the Trial Coordinator.

All SAEs (including *serious adverse device effects* or *SADEs*) must be reported to the MHRA by the sponsor or a delegated and appropriately qualified member of the research team. All SAEs which are related to the study and all SAEs which are unexpected will be reported to the REC (see protocol v9.0 for detail). SAEs or SADEs which are both related to the device and unexpected (an unanticipated serious adverse device effect – USADE) must be reported immediately and no later than 15 calendar days from Sponsor or CI awareness of the event to the Health Research Authority (HRA). All AEs/SAEs will be summarised and reported in the Open report of the Data Monitoring Committee (DMC) which will also be circulated to the Trial Steering Committee (TSC). SAEs will also be circulated to the DMC chair for review. Action will then be taken accordingly depending on implications for the conduct of the trial.

AEs and SAEs will be reported in the primary analysis in the same way as shown below. For details of the definitions of expectedness and relatedness please see the protocol. The following information will be recorded:

- Diagnosis or symptoms of adverse event
- Body system code
- Start date
- Whether the symptoms are ongoing at the end of participant follow-up or the stop date if symptoms have stopped
- Outcome of the event (recovered, recovered with sequelae, continuing, death, unknown)
- Potential relatedness to study participation,
- Whether the event is serious and if serious what type of serious event (death, life-threatening, requires hospitalisation or prolongs current hospitalisation, persistent or significant disability or incapacity, congenital anomaly or birth defect, is otherwise considered medically significant by the investigator).
- Whether the event is unexpected.

In addition to the collection, monitoring and reporting of AEs and SAEs described above, participants will be asked at 6-month follow-up to complete a self-report measure of potential adverse effects from trial involvement. Participants will rate 27 statements on a five-point scale from ‘not at all’ to ‘very much’.

1.8 Sample size estimation

We will randomise a maximum of 1064 participants (approximately 335 Northeast London; 224 for Manchester, East London and Glasgow; and 111 for Oxford); this would provide a total of 1118 participants including the 54 participants allocated to BrighterSide before removal of that arm. We power for an expected difference in proportion of admission rates between arms of 7.5%, and an expected admission rate within 6 months in the TAU group of 15% (based on extensive audit data from our NHS Trusts). The design follows the methods outlined for MAMS designs with binary outcomes calculated using - nstagebin- in Stata (Choodari-Oskooei,

2023). It uses the same outcome at both Stages and has a primary endpoint of 6 months after recruitment in each stage and the loss-to follow-up rate is 5%. One-sided significance levels of 30% and 2.5% and powers of 92% and 90% are used in the 1st and 2nd stages respectively. The total sample size of 1064 represents the maximum possible sample size under the 2-stage MAMS trial design: the minimum sample size if both remaining arms are dropped for futility at the interim analysis would be 684.

The interim analysis will take place after 385 participants provide 6-month outcome data, which will occur in May 2024. The interim analysis may result in one or both interventions being removed from the trial at stage 1 (power=92%, one-sided alpha=0.3). The final analysis has 90% power and a one-sided alpha = 0.025. The overall family-wise error rate (FWER) is 0.0397. There is no need to account for multiple testing in a multi-arm trial with independent comparators. We have allowed for a 5% missing data in the primary outcome (this is routinely available in electronic medical records, so we should have less than 5% attrition which are likely to be withdrawls)

1.9 Brief description of proposed analyses

The trial statistician will remain blind to treatment allocation until the SAP has been finalised and approved; the senior trial statistician will remain blind until all analysis has been completed. Interim analysis of the primary outcome (admissions) will be carried out by the unblinded trial statistician and checked by an independent blinded trial statistician. Final analyses will be carried out by the trial statistician and overseen by the senior trial statistician.

All analysis, both interim and final analyses will be conducted on the intention-to-treat (ITT) sample (i.e., all randomised will be analysed according to which study arm they have been allocated, regardless of the intervention or treatments actually received, adherence to treatment, or the number of measurements recorded).

For the analysis of the primary outcome at both Stages, we will use a generalised linear model that includes fixed effects for treatment, diagnosis and site using maximum likelihood estimation. For each pairwise comparison with TAU, we will report the difference in proportion of admissions as the estimate of treatment effect with 2-sided 95% confidence intervals and 1-sided p-values.

For Stage 1 interim analysis: only the primary outcome will be formally analysed. Interim analysis decision rules for the discontinuation of intervention arms to Stage 2 are described in **section 3.1**.

2 Data analysis plan – Data description

2.1 Recruitment and representativeness of recruited patients

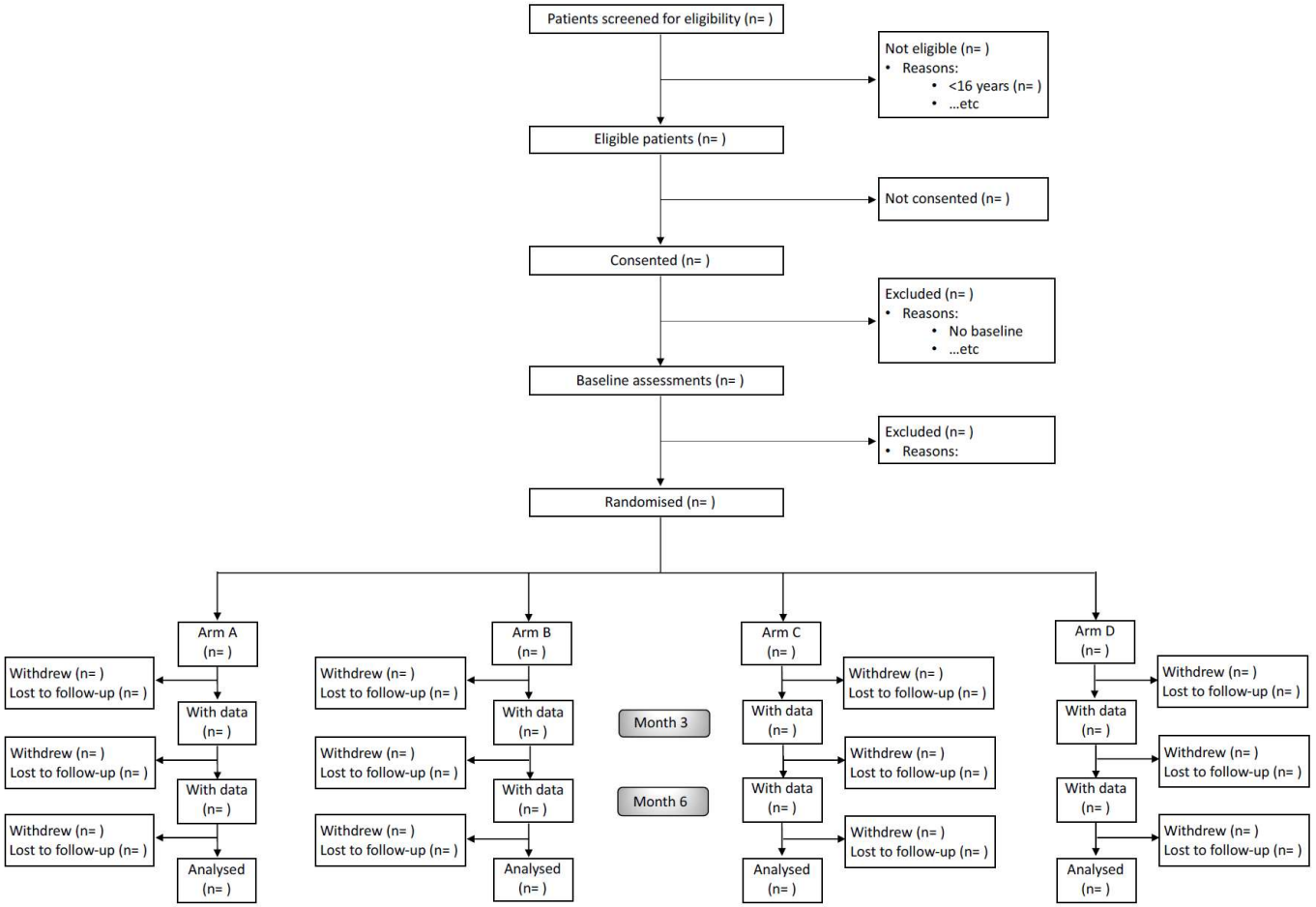
CONSORT flow chart will be constructed (Moher, 2001) – see Figure 2. This will include the number of eligible patients, number of patients agreeing to enter the trial, number of

patients refusing, then by treatment arm: the number of patients adhering/not adhering to treatment, the number continuing through the trial, the number withdrawing, the number lost to follow-up and the numbers excluded/analysed.

Adherence is defined as follows for the treatment arms:

- SAFETEL: attending the initial safety planning session and at least one follow-up call.
- PREVAIL: attending at least two peer support sessions.

2.1.1 Figure 2 - Template CONSORT diagram for RAPID trial



2.2 Baseline comparability of randomised groups

Descriptive statistics within each randomised group will be presented for baseline values. These will include counts and percentages for binary and categorical variables, and means and standard deviations, or medians with lower and upper quartiles, for continuous variables, along with minimum and maximum values and counts of missing values. There will be no tests of statistical significance or confidence intervals for differences between randomised groups on any baseline variable.

2.3 Adherence to allocated treatment and treatment fidelity

Adherent and non-adherent subgroups within the SAFETEL and PREVAIL intervention arms will be described in terms of baseline variables.

The reasons for withdrawal from treatment will be summarised.

2.4 Loss to follow-up and other missing data

The number of participants with missing data for each outcome will be summarised in each arm and at each time point. The number withdrawing from treatment and reasons for withdrawal from treatment will be summarised by treatment group and overall. For those who withdraw from treatment only, follow-up data will still be collected where possible.

The frequency and proportion of people consented who withdraw from the trial will be presented. If a participant withdraws from the trial, then no further follow up data will be collected. In particular we will check if adherence with treatment, (using adherence definitions described in section 2.1), is predictive of missingness at later time points as this may be informative as to our missing data approach.

The reasons for withdrawals and loss to follow-up will be summarised.

2.5 Adverse event reporting

Adverse events (AE), adverse reactions (AR), serious adverse events (SAE), serious adverse reactions (SAR), adverse device effects (ADE) and serious adverse device effects (SADE) will be summarised by randomised trial arm.

Participant-rated adverse effects of trial participation, assessed by the Learning from You instrument on exit from study at 6-month assessment, will also be reported. We will present proportions and frequencies of the Likert scale options chosen by the participants for the 27 items split by trial arm. Inferential analyses will not be conducted between arms for these outcomes.

2.6 Assessment of outcome measures (unblinding)

Our primary outcome is objective (admission to hospital (yes/no), so should not be subject to rater bias. Assessors (RAs) will be blind to treatment condition. All blind breaks are recorded by the Trial Manager and reviewed by the Chief Investigator for patterns in unblindings. The independent DMC and TSC will regularly monitor unblindings by each centre and implement corrective action if necessary. When a blind break does occur, where possible, we will identify an independent assessor with whom the blind has not been broken to complete subsequent follow-ups, subject to any threats to participant engagement with follow-up.

2.7 Descriptive statistics for outcome measures

The primary and secondary measures will be summarised using appropriate summary statistics, for the entire trial population and by trial group at each time point. The distributions of the continuous outcome measures will be inspected, and a judgement made on whether the variables are normally distributed or not. The mean and standard deviation will be presented for all normally distributed measures, median and quartiles for skewed distributions and proportions and frequencies for categorical measures.

The primary outcome (proportion admitted to psychiatric hospital in 6 months) and secondary outcomes proportions (for binary outcomes) and mean/medians, and confidence intervals will be plotted over time in unadjusted profile plots.

3 Data analysis plan – Stage 1 interim analysis

The interim analysis will take place after 385 participants provide 6-month outcome data, which will occur in May 2024. The interim analysis may result in one or both interventions being removed from the trial at stage 1 (power=92%, one-sided alpha=0.3).

3.1 Decision rules

The measures to safeguard the confidentiality of the interim analysis results and minimize potential operational bias are detailed below:

1. The significance level will be used as the critical value for the p-value of the observed treatment effect as this controls the type I error rate at the nominal level regardless of the true control event rate (Bratton, 2013). We designed the trial to have a non-binding futility bound with a one-sided p-value of 0.3.
2. The independent DMC will receive a closed report prepared by the unblinded trial statistician showing the between-group differences of each intervention arm compared with TAU
3. If both intervention arms meet the criteria to continue in Stage 2 (one-sided p-value <=0.3), no interim analysis results will be shared with the TMG or TSC, and the study will continue as three-arm trial, provided this is endorsed by the TSC on the recommendation of the DMC.
4. If one or both interventions arms just exceeds the threshold for futility (one-sided p-value >0.3 and <0.4), we will provide the DMC with additional information about the

between-group differences on the CSSRS to guide a decision on whether one or both arms should continue in Stage 2.

5. If one intervention arm clearly exceeds the threshold for futility (one-sided p-value >= 0.4), or if following the additional information on CSSRS the DMC recommends dropping one arm, the interim analysis results for this comparison only will shared with the TMG and TSC on the recommendation of the DMC. This arm will be dropped from Stage 2 of the study. The interim analysis results of the remaining arm will not be shared with the TMG or TSC.
6. If both arms are recommended to be dropped, following confirmation of the recommendation of the TSC by the funder and sponsor, we will terminate recruitment to the study for lack-of-benefit in either arm.

3.2 Statistical model

For Stage 1 interim analysis, only the primary outcome will be formally analysed. We will use a generalised linear model that includes fixed effects for treatment, diagnosis and site using maximum likelihood estimation. For each pairwise comparison with TAU, we will report the difference in proportion of admissions as the estimate of treatment effect with 2-sided 95% confidence intervals and 1-sided p-values.

4 Data analysis plan – Stage 2 inferential analysis

4.1 Primary analysis description

4.1.1 Definition of primary estimand

The primary clinical question of interest is: What is the difference in proportion of having at least one psychiatric hospital admission during the 6 months post-randomisation among persons with SMHPs who have had recent suicidal thoughts or behaviours (as defined by trial inclusion/exclusion criteria as per **section 1.7.1**) that are assigned to one of two brief, remote psychosocial interventions compared to TAU?

The estimand is defined using the following attributes:

- **Population**: persons with SMHPs who have had recent suicidal thoughts or behaviours meeting the trial inclusion/exclusion criteria.
- **Endpoint**: any psychiatric hospital admission between randomisation and 6 months post-randomisation.
- **Treatment condition**: TAU and SAFETEL, TAU and PREVAIL, will each be compared against TAU alone. All randomised participants will be included in the statistical analysis and analysed according to the group they were originally assigned, regardless of adherence to actual treatment.
- **Intercurrent events**: All types of intercurrent events, which include:
    - Failure to adhere to randomised treatment protocol
    - Premature discontinuation of the randomized treatment
    - Death of participant during trial period

    - Long-term non-psychiatric hospital admission

    will be handled using the treatment policy strategy, that is, included in the treatment regimen under evaluation. We will explore the pattern of withdrawals to ensure follow up data is missing at random and similar by arm. If not, remedial action will be taken by the blinded senior statistician.
- **Population-level summary**: difference in proportion of any psychiatric admission in 6 months post-randomisation between treatment conditions.

Rationale for estimand: To compare rates of preventable psychiatric hospital admissions of SAFETEL, PREVAIL against TAU.

For the primary estimand, the trial will take a pragmatic treatment policy approach to intercurrent events.

4.1.2 Further details on handling of intercurrent events

We will conduct a supplementary analysis, to assess the effect of adherence to treatment as defined in the protocol. Details of this are described in **section 4.2.8**.

4.1.3 Intention to treat population set

As per **section 4.1.1**, The main analysis will include all randomised participants (eligibility as defined in **section 1.7.1**) in the arms to which they were randomised, regardless of what treatment they received post-randomisation, with the intent to estimate an intention-to-treat (ITT) type estimand. This will be estimated for the primary outcome and all secondary outcomes.

4.1.4 Analysis of primary outcome

The primary outcome (any psychiatric admission by 6 months, yes/no) will be analysed using a generalised linear model with a binomial distribution and identity link. We will include fixed effects for randomised group, site, and indicator variables for the presence of the following SMHPs: schizophrenia spectrum disorder, bipolar disorder, major depressive disorder, or emotionally unstable personality disorder.

A pairwise comparison of all experimental arms to control arm will be made, including the arm dropped for further randomisations in Stage 1. For each pairwise comparison with TAU, we will report the difference in proportion of admissions as the estimate of treatment effect with 2-sided 95% confidence intervals and 1-sided p-values. If the treatment effect estimate on the primary outcome is significant at the α=0.025 one-sided level in the final analysis (accounting for alpha spending approach for the interim analysis) then the intervention arm will be declared superior to the TAU arm.

4.1.5 Analysis of secondary outcomes

A pairwise comparison of all experimental arms to control arm will be made for secondary outcomes as well, including the arm dropped for further randomisations in Stage 1. The α level, one-sided pairwise error rate, and FWER will be identical to the analysis of the primary outcome.

Secondary outcomes with repeated measures at 3 and 6 months (with the exception of psychiatric hospital admission) will be analysed with generalised linear mixed models appropriate for the distribution of the outcome, with a random-intercept to account for clustering of observations within participant. Secondary outcomes measured at one timepoint only will be measured generalised linear models only (i.e., no random effects).

4.1.5.1 Continuous outcomes: 3-month and 6-month repeated measures

Secondary continuous outcomes measured at 3-month and 6-month follow-up (Suicidal ideation score and intensity as measured by C-SSRS, Process of Recovery Questionnaire, ReQoL-10, GAD-7, PHQ-9, AHS, Entrapment Scale Short-Form) will be analysed with (longitudinal) generalised linear mixed regression models using both 3-month and 6-month timepoints. As with analysis of the primary outcome, fixed effects will include: treatment, site, and SMHP diagnosis, with additional adjustment for outcome measure at baseline, time (as category), and time*treatment arm interactions. The model will include a participant-level random intercept.

4.1.5.2 Binary outcomes: 3-month and 6-month repeated measures

Secondary binary outcomes measured at 3-month and 6-month follow-up (presence of suicidal ideation and occurrence of NSSI as measured by C-SSRS) will be analysed with (longitudinal) generalised logistic mixed regression models using both 3-month and 6-month timepoints. Covariate adjustment and random intercept will be as described for the secondary outcome repeated measure analysis in **section 4.1.5.1**.

4.1.5.3 Other secondary outcomes

Psychiatric hospital admission at 3 months, and suicide attempt in the last 6 months as measured by C-SSRS (both dichotomous variables) will be analysed identically to the primary outcome analysis described in **section 4.1.4**.

4.1.5.4 *Combining data across trial stages*

We will fit the analysis models separately to data from each stage of the trial. We split Stage 1 into two groups, before and after dropping Brighterside, to account for the change of randomisation ratios from 2:2:2:3 to 1:1:2.

We will use a combination test to calculate the overall p-value, combining p-values estimated separately from each of three models. We present the difference in proportions averaged across all models.

4.1.5.5 Analysis of BrighterSide arm

There were 54 participants allocated to the BrighterSide arm before recruitment to that arm was stopped. We will perform a separate analysis of this arm with the concurrently recruited controls to obtain an estimate of the effect of BrighterSide and TAU versus TAU alone on the primary outcome and the C-SSRS. We will use the same statistical models as the primary and secondary analysis outlined above, and report using the same treatment effect measures. We acknowledge these analyses will be underpowered.

4.2 Statistical considerations

4.2.1 Rationale for using maximum likelihood estimation

We are aware that the issue of point estimation and degree of bias as a result of interim stopping/selection rules is a hot topic in the statistical literature. However, the literature shows that the issue of bias is not a major concern if the design does not include interim efficacy stopping rules – which is the case for this design since arms are dropped for futility but we have no stopping rules for efficacy. Bratton et al (Bratton 2023) performed simulation studies and showed that for trials with binary outcomes the degree of bias at the final primary efficacy analysis is typically very small if using maximum likelihood estimation. In some cases, the bias can be large but the probability of those cases happening is very low.

A series of paper by Robertson and colleagues (2023a, 2023b) have addressed the issue of point estimation and bias, standard errors and confidence intervals in group sequential designs. As a result, we will apply a Uniformly Minimum Variance Conditionally Unbiased Estimator as outlined in Lee et al (2022) to compare against our results from the GLM with MLE.

4.2.2 Time points

Psychiatric hospital admission will be taken directly from electronic health records.

However, all available data will be used for the main analyses regardless of compliance to visit windows.

4.2.3 Stratification and clustering

Randomisation has been stratified by site. Therefore, it is important to include this variable in the modelling process.

The structure of secondary outcome data is longitudinal with repeated measurements. This correlation of observations within participants will be accounted for in the modelling process by including a subject-varying random intercept.

4.2.4 Missing items in scales and subscales

The number (%) with complete data will be reported. If missing data guidance is provided for missing items within the scale scoring guide, this will be used (see Appendix 1). If no guidance is provided scales will be pro-rated for an individual if 20% or fewer items are missing. As an alternative, scales will be pro-rated for an individual if 20% or fewer items are missing. For example, in a scale with 10 items, prorating will be applied to individuals with 1

or 2 items missing. The average value for the 8 or 9 complete items will be calculated for that individual and used to replace the missing values. The scale score will be calculated based on the complete values and these replacements.

4.2.5 Missing baseline data

Missing baseline data should not be an issue for the primary analysis. Some extensions to this analysis may use other baseline variables; if these contain missing data, the number with complete data will be reported and they will be imputed using a method suitable to the variable as per the recommendations of White and Thompson (2005).

4.2.6 Missing outcome data

Where there are two or more outcome time points, missing post-randomisation assessments will be dealt with by fitting linear mixed models to all the available data using maximum likelihood methods. Such an approach provides valid inferences under the assumption that the missing data mechanism is ignorable (i.e. MAR). If post treatment variables such as compliance with treatment are found to be predictive of drop out, multiple imputation will be considered.

For the primary outcome, if a participant had not had a hospitalisation at the 3-month follow up and is then lost where their EHR data cannot be obtained, their primary outcome will be set to missing.

4.2.7 Method for handling multiple comparisons

As we have a single prespecified primary outcome, all secondary analyses can be considered as exploratory therefore adjusting for multiple comparisons is not necessary. There is no need to account for multiple testing in a multi-arm trial with independent comparators (Parker & Weir, 2020). The estimated overall family-wise error rate (FWER) is 0.0403 for the trial.

4.2.8 Method for handling non-compliance (per protocol/CACE analyses)

In addition to the primary intention-to-treat analysis, the effect of adhering to treatment for each experimental arm as defined in **section 2.1** will also be estimated through a supplementary analysis.

4.2.8.1 Per-protocol analysis set

The per-protocol population (PPP) will be defined as all participants that adhered to the interventions' protocols, and participants that are classified as non-adherers will be removed from the PPP sample. Adherence is defined as follows for the treatment arms:

- SAFETEL: attending the initial safety planning session and at least one follow-up call.
- PREVAIL: attending at least two peer support sessions.

4.2.8.2 Per-protocol analysis

Analyses for the primary outcome (hospital admissions) will be repeated on the PPP sample as supplementary analyses at Stage 2 only. If proportion showing adherence for any of the interventions is found to be very low (less than 30% of those randomised to peer support arm are found to be compliant), a complier average causal effect (CACE) analysis for the primary outcome will be considered to investigate the effect of actually receiving the intervention.

4.2.8.3 Complier average causal effect analysis

If the proportion of adherence for any of the interventions is found to be lower than 80%, a complier average causal effect (CACE) analysis will be considered to investigate the effect of actually receiving the intervention on the primary outcome.

4.2.9 Model assumption checks

For the secondary outcome continuous models, one assumes normally distributed residuals. Linear mixed model residuals will be plotted to check for normality and inspected for outliers. If the distribution of the residuals deviates substantially from normal, we will fit the models requesting robust standard errors.

Any issues that arise with model assumptions that lead to a change of the planned analysis will be reviewed by the blinded senior statistician and approved by the blinded independent TSC statistician.

4.2.10 Planned subgroup analyses

There are pre-specified subgroup analysis by site and diagnosis. This will be tested by including an interaction term between the subgroup variable and treatment group in the primary analysis model.

4.3 Software

Data management: An online data collection system for clinical trials (MACRO; InferMed Ltd) will be used. This is hosted on a dedicated server at KCL and managed by the KCTU. The KCTU Data Manager will extract data periodically as needed and provide these in Stata format to the analyst(s). No other member of the trial team should be requesting data extracts, particularly any extracts containing outcome data.

Statistical analysis: Stata 18 or later will be used for data description and the main inferential analysis.

5 Reference list

Bratton D, Phillips P, Parmar M. A multi-arm multi-stage clinical trial design for binary outcomes with application to tuberculosis. BMC medical research methodology. 2013;13:139.

Choodari-Oskooei, B., D. J. Bratton, et al. (2023). "Facilities for optimising and designing multi-arm multi-stage (MAMS) randomised controlled trials with binary outcomes." Stata Journal, 23(3)

Lee KM, Robertson DS, Jaki T, Emsley R. The benefits of covariate adjustment for adaptive multi-arm designs. Stat Methods Med Res. 2022 Nov;31(11):2104-2121. doi: 10.1177/09622802221114544

Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001 Apr 14;357(9263):1191-4.

Parker RA, Weir CJ. Non-adjustment for multiple testing in multi-arm trials of distinct treatments: Rationale and justification. Clin Trials. 2020 Oct;17(5):562-566. doi: 10.1177/1740774520941419

Robertson DS, Choodari-Oskooei B, Dimairo M, Flight L, Pallmann P, Jaki T. Point estimation for adaptive trial designs I: A methodological review. Stat Med. 2023 Jan 30;42(2):122-145. doi: 10.1002/sim.9605

Robertson DS, Choodari-Oskooei B, Dimairo M, Flight L, Pallmann P, Jaki T. Point estimation for adaptive trial designs II: Practical considerations and guidance. Stat Med. 2023 Jun 30;42(14):2496-2520. doi: 10.1002/sim.9734.

White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med 2005 Apr 15;24(7):993-1007.

Appendix 1: Schedule of assessments and measures

Procedures / Measures	Consent/Recruitment	Baseline	Randomisation	3 months (end of treatment)	6 months (primary endpoint)	Ongoing
Verbal consent to contact	X					
Registration	X					
Eligibility	X					
Written consent	X					
Demographics		X				
Randomisation			X			
Psychiatric Hospital Admissions (Section A)				X	X	
Psychiatric Hospital Admissions Log (Section A)						X
Psychiatric Hospital Admissions (Section B and C)		X		X	X	
Abridged C-SSRS		X		X	X	
Process of Recovery Questionnaire (QPR)		X		X	X	
EQ-5D-5L		X		X	X	
Recovering Quality of Life (ReQoL-10)		X		X	X	
Adapted EPQ (Section A)		X		X	X	
Adapted EPQ (Section B)		X		X	X	
GAD-7		X		X	X	
PHQ-9		X		X	X	
Adult HOPE Scale (AHS)		X		X	X	
The Entrapment Short-Form Scale (E-SF)						X
Adverse Effects - Learning From You					X	
Concomitant Medications Log		X		X		
Withdrawal Form		X		X	X	
SAFETEL Adherence Log						X
PREVAIL Adherence Log						X
BrighterSide Mobile App Adherence						X
Adverse Events Log						X

Appendix 2: Scoring guidelines of clinical secondary outcome measures

Columbia Suicide Severity Rating Scale

The abridged **Columbia Suicide Severity Rating Scale** (C-SSRS) used in RAPID is an assessment tool designed to measure suicidal ideations, suicidal behaviours, and self-injurious behaviour without suicidal intent. The scale is an abridged version that uses of the original scale. The composite endpoints assessed in RAPID are **suicidal ideation**, **suicidal ideation score**, **suicidal ideation intensity, suicide attempt, and NSSI** (the latter two ask whether suicidal or NSSI behaviours occurred in six months prior to baseline and within the trial period at 6-month follow-up). The endpoints are scored according to published guidance as described below:

Suicidal ideation (binary): A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS. Score missing if one of the five questions is missing, and a participant has answered "no" to each other item.

Suicidal Ideation score: The maximum suicidal ideation category (1-5 on the CSSRS) present at the assessment. Assign a score of 0 if no ideation is present. Score missing if one of the five questions is missing, and a participant has answered "no" to each other item *more severe* than the missing item.
Suicidal ideation intensity: Add the five intensity item scores to create a total score (range 3 to 25) to represent the intensity rating. If the patient did not endorse any suicidal ideation set the intensity rating to 0. As no guidelines are provided for missing data, a conservative approach of standard prorating if no more than one missing value is present (20% of items) will be used.
Suicide attempt (binary; last six months; baseline & six-month follow-up only): Score 1 if participant reports a suicide attempt in the past six months.
NSSI (binary; last three months): Score 1 if participant reports a suicide attempt in the past six months, 0 otherwise (unless missing).

Process Recovery Questionnaire (QPR)

The **Process Recovery Questionnaire (QPR)** has 15 items each scored on a 4-point scale (0= disagree strongly, 1=disagree, 2=neither agree nor disagree, 3=agree, 4=agree strongly). Higher scores are indicative of recovery. Scores are summed to give a total.
As no guidelines are provided for missing data, a conservative approach of standard prorating if 20% or less missing values will be used.

Recovering Quality of Life (ReQoL-10)

Each question in **ReQoL-10** is scored from 'None of the time' to 'Most or all of the time'. The scores are found as a subscript under each response option. In the ReQoL-10, there are 6 positively worded questions and 4 negatively worded questions. The positively worded questions are: Q2, Q4, Q5, Q7, Q8 and Q10. They are scored from 0 to 4.

ReQoL-10 scores are calculated if no more than one item had a missing response. If a single question is unanswered, the mean value of the other responses can be imputed. If more than one question is unanswered, then the overall index score cannot be calculated. If respondents give two answers to a single question, it is recommended that the lower quality of life response is adopted.

Depression (PHQ-9)

The **PHQ-9** scores each of the nine DSM-IV criteria as "0" (not at all) to "3" (nearly every day). Symptom severity can be summarised by the total of all values. If one missing value is present (< 20% of all items), this value is imputed by the mean of the completed items (Kroenke et al. 2010).

Anxiety (GAD-7)

The **GAD-7** score 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, and adding together the scores for the seven questions. If one missing value is present (< 20% of all items), this value is imputed by the mean of the completed items (Kroenke et al. 2010).

Adult HOPE Scale (AHS)

AHS is a 12-item measure of a respondent's level of hope. In particular the scale is divided into two subscales that comprise Snyder's cognitive model of hope: (1) Agency (i.e., goal-directed energy) and (2) Pathways (i.e., planning to accomplish goals). Of the 12 items, 4 make up the Agency subscale and 4 make up the Pathways subscale. The agency subscale score is derived by summing items 2, 9, 10, and 12; the pathway subscale score is derived by adding items 1, 4, 6, and 8. The total Hope Scale score is derived by summing the four agency and the four pathway items.
As no guidelines are provided for missing data, a conservative approach of standard prorating if 20% or less missing values will be used (no more than one item).

Entrapment Scale Short Form (4-item scale)

The **Entrapment Scale Short Form** has five ordered response options (0 = "not at all like me", 1 = "a bit like me", 2 = "moderately like me", 3= "quite a bit like me", 4= "extremely like me"). The secondary outcome is the sum score of all items. No pro rating will be used for missing items.

Adverse Effects – Learning From You

The **Learning From You** instrument scores each of the 27 items as "1" (not at all) to "5" (nearly every day). Mean scores of each item will be summarised by item for each arm of the study descriptively only.

EQ-5D-5L

Original Paper

Janssen, M.F., Pickard, A.S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., Swinburn, P. and Busschbach, J., 2013. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of Life Research*, *22*(7), pp.1717-1727.

Scoring Guidance

The EQ-5D-5L descriptive system comprises five dimensions (MOBILITY, SELF-CARE, USUAL ACTIVITIES, PAIN / DISCOMFORT and ANXIETY / DEPRESSION). Each dimension has five response levels: no problems, slight problems, moderate problems, severe problems, unable to/extreme problems. Responses are coded as single-digit numbers expressing the severity level selected in each dimension. For instance, 'slight problems' (e.g. 'I have slight problems in walking about') is always coded as '2'. The digits for the five dimensions can be combined in a 5-digit code that describes the respondent's health state; for instance, 21111 means slight problems in the mobility dimension and no problems in any of the other dimensions. They have no arithmetic properties. For instance, on the basis of just the numbers one cannot assume that a state 21111 is better 13111. Therefore, these numbers should not be used to derive a summary score. To derive the summary indexscore, an appropriate 'value set' is required.

Health economist to input – will not be analysed by trial statisticians

Appendix 3: Summary of blinding/unblinding

Study Role	Blinding Status (participant level)	Planned unblinding (participant level)	Blinding Status (summary level)	Planned unblinding (summary level)
Chief Investigator	Blinded	Unblinded if required for SAE/AE evaluation	Blinded	Unblinded upon review of final study report
Senior Statistician	Blinded	Unblinded upon review of final study report	Blinded	Unblinded upon review of final study report
Junior Statistician	Unblinded	Unblinded after preparing first DMEC closed report	Unblinded	Unblinded after preparing first DMEC closed report
Outcome Assessor	Blinded	Once participant completes trial	Blinded	Unblinded upon review of final study report
Trial co-ordinators	Unblinded	Responsible for randomising participants, will be unblinded upon randomisation of participants	Blinded	Unblinded upon review of final study report
Independent Data Monitoring Committee members	Blinded	Unblinded upon review of final study report	Unblinded	From first DMEC closed sessions where safety and outcome data is reviewed by arm
Independent TSC members	Blinded	Unblinded upon review of final study report	Blinded	Unblinded upon review of final study report

If there are instances of unplanned unblinding, these will also be recorded in the study unblinding log maintained by the trial co-ordinators.