Statistical Analysis Plan

The clinical and cost effectiveness of behavioural therapy for interepisode bipolar symptoms (STABILISE): A feasibility study

Protocol Version and Date:	V1.3 24 OCT 2024
Trial Registration No:	ISRCTN 18207465
IRAS ID or other identifier:	NIHR302220
Chief Investigator:	Prof Kim Wright
Senior Trial Statistician:	Assoc Prof Fiona Warren
SAP Author:	Assoc Prof Fiona Warren

Version History:

Version and date (dd/mmm/yyyy)	Description of changes	Stage of Revision e.g., before interim analyses	Justification of Revision
1.0 (12/FEB?2025)	N/A	Prior to analysis	N/A

Roles and Responsibilities:

Approved by:

Senior Trial Statistician

Signature: DocuSigned by: Fiona Warren 7E787DA05B58417... Date: 12 February 2025 | 12:02 GMT (dd/mmm/yyyy)

Name (please print): Assoc Prof Fiona Warren

Chief Investigator

Signature: DocuSigned by: Kim Wright 06EEF378E1FC4BC... Date: 12 February 2025 | 12:54 GMT (dd/mmm/yyyy)

Name (please print): Prof Kim Wright

Independent Statistician

Signature: DocuSigned by: Yvonne Sylvestre CDE7CB79A38F469... Date: 14 February 2025 | 08:37 GMT (dd/mmm/yyyy)

Name (please print): Dr Yvonne Sylvestre

Table of Contents

1. Introduction 4
1.1. Background and rationale 4
1.2. Objectives 4
2. Study Methods 5
2.1. Trial Design 5
2.2. Randomisation 6
2.3. Sample Size 7
2.4. Framework 7
2.5. Statistical Interim analyses and stopping guidance 7
2.6. Timing of final analysis 7
2.7. Timing of outcome assessments 7
3. Statistical Principles 9
3.1. Confidence intervals and p values 9
3.2. Intervention adherence and protocol deviations 9
3.3. Analysis populations 10
4. Trial population 10
4.1. Screening data 10
4.2. Eligibility 10
4.3. Recruitment 11
4.4. Withdrawal / follow-up 12
4.5. Baseline patient characteristics 13
5. Statistical Analysis 13
5.1. Outcome definitions 13
5.2. Analysis Methods 15
5.3. Missing data 16
5.4. Additional analyses and sensitivity analyses 16
5.5. Harms 16
5.6. Statistical software 16
6. Related Documents 16
7. References 17
8. Appendices 19

1. Introduction

1.1. Background and rationale

Bipolar Spectrum Disorders typically involve periods of depression, and periods of very high energy and mood (hypomania or mania). Bipolar Disorders (BDs) are common, affecting around 1 in 20 people at some point in their lifetime. They can be very distressing and disruptive to the people who experience them, and for friends and relatives. Many people with these conditions have mood issues outside depressive or manic episodes. Often these issues include ongoing low mood, and/or frequent swings in mood or emotions. These are sometimes called “inter-episode bipolar symptoms” (IEBS).

There are psychological (talking) therapies designed to help people with BDs, but these tend to be aimed at preventing a relapse of depression or mania, or they are aimed at helping people recover from a period of depression. There is no universally accepted talking therapy for helping people who have ongoing low mood or mood swings in between full episodes. Hence, further research is needed to develop therapies that would support people with BDs who experience IEBS.

1.2. Objectives

The aim of this feasibility trial is to inform the design and conduct of a future definitive trial to investigate a psychological therapy designed to address these ongoing symptoms.

The specific feasibility objectives of the trial are as follows.

i. To inform the recruitment and timeline of a future definitive trial, by establishing the number of participants identified, approached, consented, randomised and completed.
ii. To refine future definitive trial procedures by establishing the acceptability and experience of the trial process to participants, including randomisation and completion of outcome measures.
iii. To determine the optimal primary outcome measure in a future definitive trial by assessing the performance of four selected candidate primary outcome measures with respect to level of acceptability to participants and participant-perceived relevance and value. Although not powered to detect a significant between group mean difference, we will scrutinise between group mean differences for each outcome measure, as well as relative rates of reliable improvement and deterioration, to assess sensitivity of the measure to the impact of the intervention.
iv. To inform estimation of sample size for a future definitive trial by measuring data completeness at follow up (participant attrition) and standard deviation of the likely primary outcome measure (to compare with reports in published literature).
v. To characterise the comparator condition, treatment as usual (TAU), across individuals and sites.

vi. To further assess the safety and acceptability of the treatment and, based on input from trial participants and clinicians, to further refine and develop the treatment manual and the procedures for training, supervising and assessing the competence of trial therapists.
vii. To demonstrate feasibility outside of the lead site by including a second site.
viii. To estimate the cost of the intervention.
ix. To assess the feasibility of collecting health care resource utilisation data and health-related quality of life data using EQ-5D-5L.To explore the feasibility of collecting momentary assessment (experience sampling) data at 3 time points (5 times per day, for 10 days, at each timepoint): the ultimate purpose of collecting these data in a future larger trial is to estimate mood instability and to examine process of change.

2. Study Methods

2.1. Trial Design

The STABILISE trial is a feasibility trial with two parallel groups. Participants will be allocated in a 1:1 ratio to either the active intervention group (TAU + STABILISE programme) or to the control group (TAU only).

The intervention consists of up to 20 individual therapy sessions (plus up to 2 initial assessment sessions) of behavioural therapy, delivered over up to 7 months (30 weeks). Therapy ends after 20 sessions or 7 months, whichever is sooner, and participants can choose to space sessions out over more than a week if they wish or to take a short break from therapy. Sessions are 60 minutes long as default but with the option for participants to agree shorter or longer (up to 75 minute) session duration if needed (approach devised in consultation with PPI panel and therapists experienced in working with this patient group). This is followed by a period of consolidation whereby patients can opt to see the therapist up to 3 times up until 12 months after starting therapy. Therapy can be delivered face-to-face, online or by phone according to participant preference and practical constraints. Face-to-face therapy can take place in a treatment centre, in the participant's home or at an alternative venue.

Outcome measures will be recorded at baseline (pre-randomisation) and at 3 follow-up points: 14, 30 and 52 weeks post randomisation, with 30 weeks as the primary end point (as all participants receiving TAU + STABILISE will have completed treatment by this point).
The below criteria are required to be met prior to progression to a definitive trial.

Table 1 Criteria for progression to a definitive trial

	Red	Amber	Green

Safety of intervention and trial procedures as agreed by TSC and sponsor	Cannot be made sufficiently safe for use in a definitive trial	Can be made sufficiently safe following modification	Sufficiently safe for use in a definitive trial
Mean recruitment rate per site over 15 months	<2 per month at both sites	<2 per month at one site and target sample size not reached	≥ 2 per month at both sites OR <2 per month at one site but overall target sample size reached
Outcome measure completion at 30-week follow-up point (completion of all four candidate primary outcome measures at primary end point, i.e. providing sufficient data for calculation of overall score)	<55% of participants	Between ≥ 55 and <75% of participants, with clear plan for improvement	≥ 75% of participants
Level of completion of intervention (defined as attending at least 6 therapy sessions)	<50% of participants	50-89% of participants, with clear plan for improvement	≥ 90% of participants

Red: Do not progress to the main trial.

Amber: Progress if action plan to mitigate problems can be determined and agreed with the Trial Steering Committee.

Green: Progress directly to the main trial.

2.2. Randomisation

Following completion of baseline assessments and provision of informed consent, eligible participants will be randomised on a 1:1 ratio to TAU (control group) or TAU plus STABILISE programme (TAU + STABILISE; intervention group). The randomisation process will use minimisation with two factors: (i) trial site (Devon, Avon and Wiltshire) and (ii) medication for BDs (current use, no current use). The first six participants will be allocated using simple randomisation and then the minimisation procedure will take over, maintaining a stochastic element to the algorithm to allow concealment to be maintained. The minimisation algorithm will take account of the allocations of the first six participants.

To ensure allocation concealment, eligible participants will be randomised via a validated password protected website hosted by Exeter Clinical Trials Unit. Once a participant has attended a baseline assessment and has consented to participation in the trial, an unblinded member of the research team will enter a set of data on the randomisation website (including Trial ID, site, medication status, date of consent) and will receive the participant's allocation.

Further details on the randomisation procedure are documented in the trial's Randomisation Statistical Requirements document.

2.3. Sample Size

As this is a feasibility trial, no formal sample size calculation is required. A total of 60 participants will be recruited (target n=30 in Devon, and n=30 in Avon & Wiltshire). To justify this choice of sample size, 60 participants allows estimation of loss to follow-up within +/- 12 percentage points, assuming 20% attrition, and is sufficient to estimate the standard deviation in the candidate primary outcome measures [1-2].

2.4. Framework

As this is a feasibility trial, and is not powered for any formal hypothesis testing, there is no specific hypothesis testing framework.

2.5. Statistical Interim analyses and stopping guidance

No interim analyses are planned for this feasibility trial. It is intended to continue to recruitment of 60 participants in total if achievable within time constraints. Early stopping would be considered only in the event of safety or well-being concerns for participants, and would be discussed with the Trial Steering Committee as required.

2.6. Timing of final analysis

Following completion of the final participant's 52-week follow-up, the database will be locked. Final datasets, including datasets from different sources, will be merged using unique participant identifiers. The final dataset, including treatment group allocations, will be provided to the unblinded senior trial statistician, who will then ensure that the database is blinded prior to sharing with the blinded trial statistician for analysis.

2.7. Timing of outcome assessments

Demographic and other participant characteristics will be measured at baseline.

1. Age
2. Gender
3. Ethnicity
4. Relationship status
5. Highest level of education
6. Perceived financial status
7. Currently open to secondary care services

Clinical outcome measures will be completed at baseline and at 14-, 30- and 52-week follow-up.

1. PHQ-9: Patient Health Questionnaire 9-item
2. ALS: Affective Lability Scales – short version
3. BRMS: Bech-Rafaelson Mania Scale
4. QoLBD: Brief Quality of Life in Bipolar Disorder
5. GAD-7: Generalised Anxiety Disorder 7-item
6. BRQ: Bipolar Recovery Questionnaire (BRQ)
7. Life chart self-report – self report of number and duration of depressive and manic episodes in the period since last assessment (completed at 14-, 30- and 52-week follow-up points only)
8. HEQ: Health Economic Questionnaire
9. EQ-5D-5L: European Quality of Life- 5 Dimensions – 5 Levels

Clinical outcomes that may potentially be selected as the primary outcome for a definitive trial are PHQ-9, ALS, QoLBD and BRQ.

Process measures will be completed at baseline, 14- and 30-week follow-up.

1. PU: Positive Urgency Scale
2. NU: Negative Urgency Scale
3. BADS-SF: Behavioural Activation in Depression Scale – Short Form
4. MAB: Momentary Assessment Block (current mood and activity reported 5 times a day for 10 days at each timepoint via a purpose-built web application)

Weekly assessments will be completed prior to each therapy session by those in the intervention arm only, as part of the therapy process.

1. ASRM: Altman Scale for Rating Mania

Additional measures include the following.

1. Feedback questionnaire (30-week follow-up) including the following items:
    i. Satisfaction with therapy (intervention group only)
    ii. Acceptability of therapy (intervention group only)
    iii. Would recommend to a friend (intervention group only)
    iv. Overall satisfaction with research elements of therapy (both groups)
2. Adverse events form: completed at intake assessment and at 14-, 30- and 52-week follow-up
3. Ranking of outcome measures according to which they most wish to experience change in: baseline and 30-week follow-up
4. BARS: Brief Adherence Rating Scale (adherence to prescribed medication using a single overall compliance question): intake, 14- and 30-week follow-up

The feedback questionnaire also includes the opportunity to provide written feedback on any of their responses to the questions, plus any general comments on any aspect of the therapy or research. Participants who exit the trial early will also be given the opportunity to complete a feedback questionnaire if they have indicated that they are willing to do so.

For all outcome measures at each timepoint, data collection will be up to one month following the scheduled date of collection. The exceptions to this are BARS, BMRS and SCID, which will be collected up to 2 weeks prior to 4 weeks post the scheduled date of completion. For reporting of outcomes, all data will be included, regardless of whether the data was collected within the designated window. In addition, the number and percentage of participants for whom the first presented outcome measure fell outside the window will be reported, in addition to the mean (SD) and median (IQR, min, max) for duration of time from scheduled date of collection, for each group separately. (It is noted that participants who complete outcomes online can complete the outcomes across more than one day; the date of completion of the first presented outcome measure will be taken. For participants who complete the outcomes on paper, the reported date of completion will be used; if this is missing, the date of receipt of the form will be used.) Data on timing of completion of outcomes will be used to inform the data collection windows for a definitive trial.

3. Statistical Principles

3.1. Confidence intervals and p values

Two-sided confidence intervals (CIs) for all metrics where an uncertainty estimate is required (e.g. between group mean differences, risk differences, odds ratios) will be reported at the 95% level. As this is a feasibility trial, no hypothesis testing of between group comparisons will be performed and hence no p-values will be reported for any inferential analyses.

3.2. Intervention adherence and protocol deviations

For the STABILISE+TAU group, participant adherence to the intervention will be reported descriptively in terms of number of the three types of therapy session attended:

1. Main therapy sessions (maximum of 22)
2. Consolidation sessions (maximum of 3)

Adherence of the therapists to the intrinsic content and therapeutic elements of the STABILISE programme will be measured by development of a therapy adherence and competence measure (discussed further in Protocol V1.2, 30 Aug 2024, Section 3).

For the TAU group, participants are assumed to be adherent to their allocated treatment by default, i.e. adherence to TAU cannot be measured as such. However, we will report whether any participants allocated to TAU accessed the STABILISE intervention due to error. Also, TAU will be described based on the Health Economic Questionnaire.

3.3. Analysis populations

All analyses of relevant feasibility outcomes (retention, completion of potential clinical primary outcomes and momentary assessment data, and adherence to STABILISE intervention), and all clinical outcome measures of the trial will be reported using the intention to treat (ITT) principle: participants will be analysed according to their randomised allocation, irrespective of treatment actually received. Participants allocated to STABILISE+TAU may not receive their allocated treatment at all, if they do not attend any of the therapy sessions, or may receive a dose that is so minimal as to be unlikely to have any therapeutic benefit. Participants allocated to TAU are unable to access the STABILISE programme by other means (unless in error), although it is possible they may seek support for their BD elsewhere.

Harms will be reporting using a safety population approach, i.e. participants will be classed according to whether they received at least one therapy session.

4. Trial population

4.1. Screening data

Potentially eligible individuals will be screened for eligibility prior to being requested to provide informed consent and progression to collection of further demographic data at baseline, and then to randomisation. Data on age, gender and ethnicity will be collected at the screening stage.

4.2. Eligibility

Participants aged 18 or over and currently experiencing interepisode symptoms (within the context of Bipolar I or II Disorder, other specified bipolar disorder or cyclothymic disorder) as the primary presenting problem will be eligible for recruitment into the study.

Inclusion criteria

Participants will be adults who:

i. meet research diagnostic criteria for Bipolar I or II Disorder, Other Specified Bipolar Disorder or Cyclothymic Disorder
ii. do not meet criteria for a manic or severe depressive episode
iii. have IEBS, defined as at least mild depressive symptoms (Patient Health Questionnaire-9 [PHQ-9] >=5) or above-average bipolar mood instability defined as >=1.3 on the brief Affective Lability Scale (ALS) depression-elation scale
iv. are willing to engage in psychological work addressing IEBS or its impact on functioning
v. have sufficient English proficiency to complete questionnaires without translation
vi. have completed the intake measures
vii. are registered with a General Practice within the study site catchment area.

Exclusion criteria

Exclusion criteria include:

i. current substance dependence according to ICD-11 criteria (as this may interfere with ability to engage in and use therapy; current substance abuse is not an exclusion criterion)
ii. risk of harm to self or others that cannot be safely managed in a community outpatient setting
iii. currently engaged in another psychological therapy for bipolar disorder
iv. participant anticipates they will be unable to regularly attend therapy sessions within the site area (e.g. planning to move out of the study area, work commitments prevent regular attendance, lengthy period of travel not mitigated by access to online therapy sessions).

The following measures will be completed at intake to assess participant eligibility.

1. Structured clinical interview for DSM-V (SCID-5) [3] – at baseline only, a standardised interview to establish whether the participant meets research diagnostic criteria for lifetime bipolar I or II disorder, cyclothymic disorder, other specified bipolar disorder, current depressive episode and current manic episode in order to assess whether they meet inclusion/exclusion criteria.
2. The Hamilton Depression Rating Scale (HAM-D) [4] will be used with participants who meet criteria for a current depressive episode to establish severity (those scoring in the severe range of >=24 will not be eligible).
3. The International Classification of Diseases, 11th revision (ICD-11) [5] will be used to determine presence or absence of current substance dependence. This is because DSM-V does not distinguish between substance abuse versus dependence (discriminating on the basis of severity instead), yet the former distinction is the most relevant to the patient group targeted by this intervention. The SCID, rather than review of patient records, will be used to ascertain research diagnosis for the purposes of this study.
4. PHQ-9 [6]
5. ALS [7]

In order to be eligible participants must score either >=5 on the PHQ-9 or >10 on the depression-elation subscale of the ALS.

4.3. Recruitment

The CONSORT flow diagram will include data relating to the following stages in the (potential) participant's journey through the trial.

1. Number of potential participants completing screening telephone call (with non-progression numbers due to ineligibility or choice)
2. Number of potential participants completing intake assessment (with non-progression numbers due to ineligibility or choice)
3. Number of participants randomised to each treatment group
4. Number of participants followed-up at 14 weeks (any contact with trial team, irrespective of level of data completion; full data collection), and numbers withdrawn or lost to follow-up for now (passive non-contact)
5. Number of participants followed-up at 30 weeks (any contact with trial team, irrespective of level of data completion; full data collection), and numbers withdrawn or lost to follow-up for now (passive non-contact)
6. Number of participants followed-up at 52 weeks (any contact with trial team, irrespective of level of data completion; full data collection), and numbers withdrawn or lost to follow-up for now (passive non-contact)

More detailed information on reasons for non-progression or withdrawal will be provided in a separate table.

An outline of the CONSORT flow diagram is to provide in the Appendix (Figure 1).

4.4. Withdrawal / follow-up

Should a participant be found to be ineligible for the trial post randomisation, it will be considered whether continued participation in the intervention (if in the STABILISE+TAU group) is in the best interests of the participant. If this is not the case, the participant will be withdrawn from treatment. Otherwise, the participant will continue to participate in the treatment; in both cases the participant will be requested to provide further follow-up data if it is safe and the participant is happy to do so. However, only participants who are eligible for the trial will be included in the final analyses.

If a participant withdraws consent from the trial, no further follow-up data will be sought. However, the participant's baseline data and any follow-up data already collected will be retained and included in the analyses.

If a participant is passively lost to follow-up during the trial (i.e. not responding to attempts to make contact for follow-up assessments), then any data previously collected will be included in the analyses.

If a participant in the STABILISE+TAU group wishes to discontinue treatment (i.e. not to attend any further therapy sessions), the participant will be invited to continue to provide follow-up data, and if the participant agrees, their data will be included in analyses. All of the participant's data up to the

point of treatment discontinuation will be retained for use in the analysis, even if the participant declines to provide further follow-up data.

If it is judged by a member of the trial team that discontinuation of the intervention is in the best interests of the participant, the participant will not receive any further treatment if in the STABILISE+TAU group; however, the participant will be invited to provide further follow-up data if it is safe to do so. If it considered that discontinuation in all aspects of the trial, including further collection of follow-up data, is in the best interests of the participant, then the participant will be withdrawn from further data collection, although data previously collected will be included in the analyses. If a participant dies during the trial period, then any data previously collected will be included in the analyses.

4.5. Baseline patient characteristics

Demographic and other participant characteristics will be measured at baseline.

1. Age
2. Gender
3. Ethnicity
4. Sexual orientation
5. Household members
6. Employment status
7. Perceived financial status
8. Currently open to secondary care services (Yes/No)

5. Statistical Analysis

5.1. Outcome definitions

As this is a feasibility trial, there are no primary and secondary outcomes. Outcome measures are divided into the following categories.

1. Feasibility outcome measures
2. Clinical outcome measures
3. Process outcome measures
4. Weekly assessment measures (STABILISE+TAU group only)
5. Additional measures

Feasibility outcome measures include the following.

1. Recruitment rate (overall and by site)
2. Retention at 14, 30 and 52 weeks: any contact with trial team; completion of first presented potential clinical primary outcome measure; completion of all four potential clinical outcome measures
3. Completion of momentary assessment data at 14 and 30 weeks
4. Adherence to STABILISE intervention (STABILISE+TAU group only): attendance at a minimum of six STABILISE sessions

There are nine clinical outcome measures, as follows.

1. PHQ-9: Patient Health Questionnaire-9 [6]
2. ALS: Affective Lability Scales – short version [7]
3. BRMS: Bech-Rafaelson Mania Scale [8]
4. QoLBD: Brief Quality of Life in Bipolar Disorder [9]
5. GAD-7: Generalised Anxiety Disorder-7 [10]
6. BRQ: Bipolar Recovery Questionnaire (BRQ) [11]
7. Life chart self-report – self report of number and duration of depressive and manic episodes in the period since last assessment
8. HEQ: Health Economic Questionnaire [12]
9. EQ-5D-5L: European Quality of Life- 5 Dimensions – 5 Levels [13]

All of these measures are collected at baseline, 14-, 30- and 52-week follow-up, with the exception of Life chart self-report, which is not collected at baseline.

There are three process outcome measures, as follows.

1. PU & NU: Positive Urgency and Negative Urgency Scale [14]
2. BADS-SF: Behavioural Activation in Depression Scale – Short Form [15]
3. MAB: Momentary Assessment Block (current mood and activity reported 5 times a day for 10 days at each timepoint via a purpose-built web application)

There are two weekly outcome measures (STABILISE + TAU group only).

1. ASRM: Altman Scale for Rating Mania [16]
2. BDI: Beck Depression Inventory [17]

There are four additional measures, as follows.

1. Feedback questionnaire (30-week follow-up) including the following items:

i. Satisfaction with therapy (intervention group only)
ii. Acceptability of therapy (intervention group only)
iii. Would recommend to a friend (intervention group only)
iv. Overall satisfaction with research elements of therapy (both groups)

2. Adverse events form: completed at intake assessment and at 14-, 30- and 52-week follow-up
3. Ranking of outcome measures according to which they most wish to experience change in: baseline and 30-week follow-up
4. BARS: Brief Adherence Rating Scale (adherence to prescribed medication using a single overall compliance question) [18]: intake, 14- and 30-week follow-up

Table A.1 provides details of methods of scoring for the above measures (with the exception of HEQ and EQ-5D-5L, discussed in the Health Economic Analysis Plan), for example to calculate overall scores.

5.2. Analysis Methods

All analyses will be based on the intention to treat (ITT) principle (including analyses of feasibility outcomes where appropriate): participants will be analysed according to their allocated treatment group, irrespective of treatment actually received.

All clinical and process measure outcomes are reported across both treatment groups. All outcomes will be reported descriptively at all timepoints for each group. In addition, a between group mean difference with 95% confidence interval (CI) will be reported for the primary timepoint at 30 weeks. The between group mean difference will be derived from a linear regression model, which will include treatment group as a fixed effect and will also adjust for study site and use of medication for BDs (factors used in the randomisation procedure).

Weekly outcome measures are being reported as part of the therapy process, and will not be analysed as part of the trial results.

The feedback questionnaire will be reported descriptively for those items that relate to the STABILISE+TAU group only; the overall satisfaction item will be reported descriptively by treatment group and with a between group mean difference and 95% CI (using a linear regression model as for the clinical outcomes).

Clinical outcome measures will be ranked according to preference of experiencing positive change in relation to the outcome. These ranks will be summarised using the mean rank to produce a hierarchy of preference of experiencing positive change among the outcomes.

The BARS [18] score will be reported descriptively at each timepoint by treatment group, with a between group mean difference with 95% CI reported at 30-week follow-up, using the same methods as for clinical outcomes.

5.3. Missing data

Every attempt will be made to ensure that data is collected for all outcomes, and at all timepoints where outcome data is scheduled to be collected.

For questionnaire outcomes, in the event of a pre-specified small number of missing individual items, an overall score may be calculated according to pre-defined guidelines.

Missing baseline and follow-up data will not be imputed; all analyses will include observed data only. The number of participants with data included in each analysis will be reported.

5.4. Additional analyses and sensitivity analyses

Due to the feasibility nature of this trial, no additional analyses are planned in addition to the ITT analyses including observed data only.

5.5. Harms

Adverse events data will be collected at intake, and at 14-, 30- and 52-week follow-up (also reported on an ad hoc basis during the trial). Non-serious adverse events will not be reported as part of the trial reporting. Serious Adverse Events (SAEs) will be reported as number of SAEs using the 'as treated' approach: participants in the STABILISE+TAU group will be classed as TAU only if they do not attend at least one therapy session. (It is assumed that participants allocated to TAU only will not be able to access therapy sessions.) SAEs will be reported by 'as treated' group during the 52-week trial period, and as the number of individual participants experiencing one or more SAEs during the 52-week trial period. In addition, a risk difference and odds ratio for experiencing one or more SAEs during the 52-week trial period, with 95% CIs.

5.6. Statistical software

Stata v18.0 or later, or R version 4.2 or later, will be used for the analysis.

6. Related Documents

a. Randomisation Statistical Requirements document

7. References

1. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry. 2005;4(4):287-91.
2. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Statistical methods in medical research. 2016;25(3):1057-73.
3. First MB, Williams JB, Karg RS, Spitzer RL. Structured Clinical Interview for DSM-5 Disorders–Research Version (SCID-5-RV). Arlington: American Psychiatric Association. 2014.
4. Hamilton M. A rating scale for depression. Journal of neurology, neurosurgery, and psychiatry. 1960;23(1):56.
5. World Health Organization. *ICD-11: International classification of diseases* (11th revision). 2022.
6. Kroenke K, Spitzer RL, Williams JB. The PHQ- 9: validity of a brief depression severity measure. Journal of general internal medicine. 2001;16(9):606-13.
7. Oliver MN, Simons JS. The affective lability scales: Development of a short-form measure. Personality and individual differences. 2004;37(6):1279-88.
8. Bech P, Rafaelsen OJ, Kramp P, Bolwig TG. The mania rating scale: scale construction and inter-observer agreement. Neuropharmacology. 1978;17(6):430-1.
9. Michalak EE, Murray G, CREST. BD. Development of the QoL. BD: a disorder- specific scale to assess quality of life in bipolar disorder. Bipolar disorders. 2010;12(7):727-40.
10. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Archives of internal medicine. 2006;166(10):1092-7.
11. Jones S, Mulligan LD, Higginson S, Dunn G, Morrison AP. The bipolar recovery questionnaire: psychometric properties of a quantitative measure of recovery experiences in bipolar disorder. Journal of affective disorders. 2013;147(1):34-43.
12. Simon, J, Mayer, S. HEQ (Health Economics Questionnaire), Version 08-09-2016. Vienna: Department of Health Economics, Center for Public Health, Medical University of Vienna. 2016.
13. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research. 2011;20:1727-36.

14. Cyders MA, Smith GT, Spillane NS, Fischer S, Annus AM, Peterson C. Integration of impulsivity and positive mood to predict risky behavior: development and validation of a measure of positive urgency. Psychological assessment. 2007;19(1):107.
15. Manos RC, Kanter JW, Luo W. The behavioral activation for depression scale–short form: development and validation. Behavior therapy. 2011;42(4):726-39.
16. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman self-rating mania scale. Biological psychiatry. 1997;42(10):948-55.
17. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Archives of general psychiatry. 1961;4:561-71.
18. Byerly MJ, Nakonezny PA, Rush AJ. The Brief Adherence Rating Scale (BARS) validated against electronic monitoring in assessing the antipsychotic medication adherence of outpatients with schizophrenia and schizoaffective disorder. Schizophrenia research. 2008 ;100(1-3):60-9.

8. Appendices

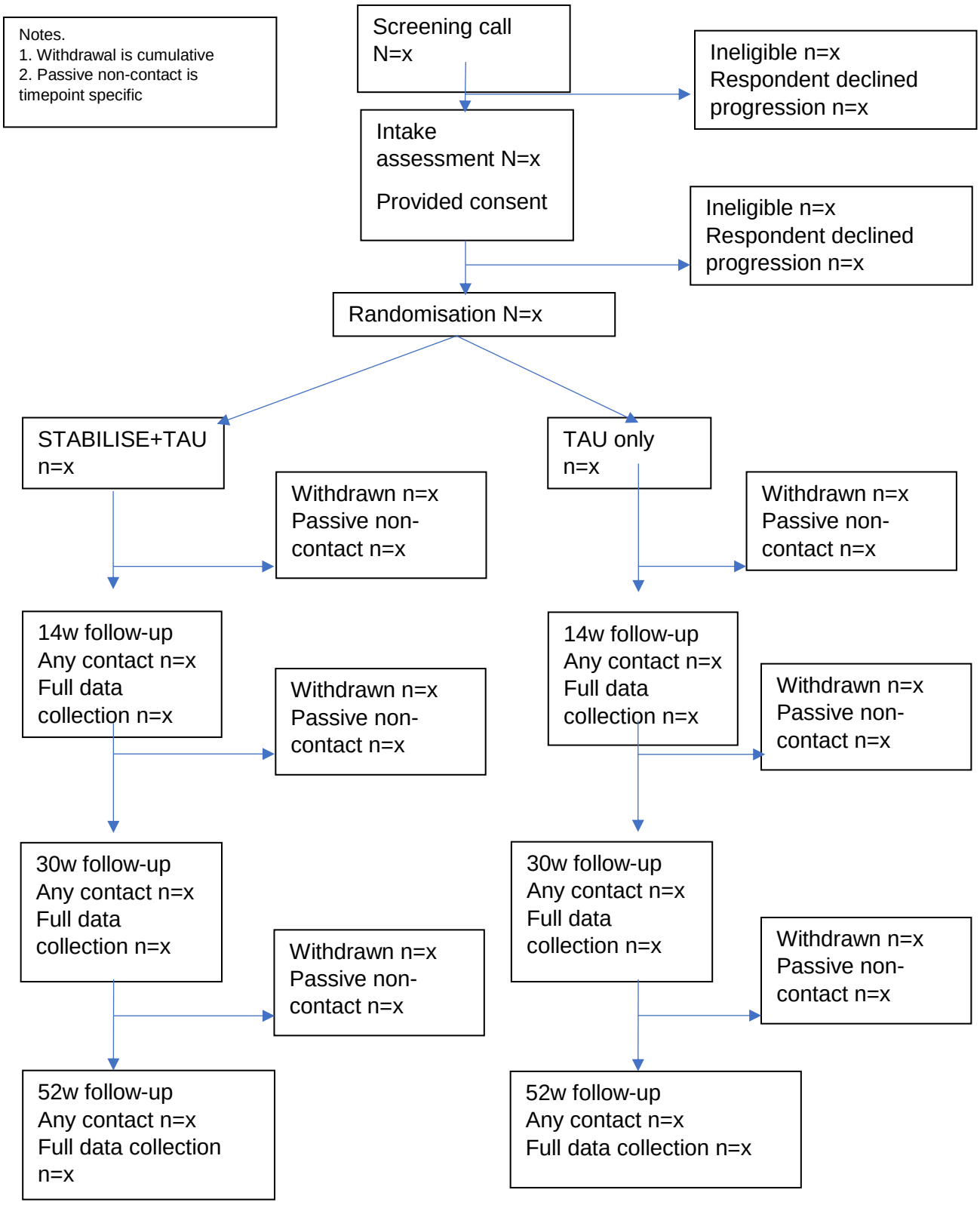

Figure A.1 Participant flow through trial

Table A.1 Calculation of overall scores for questionnaire outcomes

Questionnaire name	Acronym	Measured construct	N items	Maximum number of missing items for calculation of overall score	Subscales (N) Or Items in subscale	Individual item score (range)	Method of calculation of overall score	Overall score (range)
Patient Health Questionnaire-9	PHQ-9	Depression	9	Need all items	N/A	0-3	Sum	0-27
Generalised Anxiety Disorder-7	GAD-7	Anxiety	7	Need all items	N/A	0-3	Sum	0-21
Affective Lability Scales short version	ALS Total	Affective lability	18	Allow one missing item missing	Three	0-3	Sum	0-54
ALS Anxiety Depression	ALS A/D	Anxiety/Depression	5	Need all items	1, 3, 5, 6, 7,	0-3	Sum	0-15
ALS Depression Elation	ALS D/E	Depression/Elation	8	Need all items	2, 10, 12, 13, 15, 16, 17, 18	0-3	Sum	0-24
ALS Anger	ALS A	Anger	5	Need all items	4, 8, 9, 11, 14	0-3	Sum	0-15
Bech-Rafaelson Mania Scale	BRMS	Mania	11	Allow one missing item	N/A	0-4	Sum	0-44
Brief QoL in BD	QoLBD	BD specific QoL	12	Allow one missing item	N/A	1-5	Sum	5-60
Bipolar Recovery Questionnaire	BRQ	Sense of personal recovery	36	Allow up to three missing items	N/A	0-100	Sum	0-3600

Questionnaire name	Acronym	Measured construct	N items	Maximum number of missing items for calculation of overall score	Subscales (N) Or Items in subscale	Individual item score (range)	Method of calculation of overall score	Overall score (range)
Life Chart Self-report	N/A	Number of episodes and days in depression/ mania/hypomania	N/A	Need both number of episodes and number of days in each episode to calculate total number of days in state	Six separate variables: number of episodes and total number of days in each state	N/A	Total number of days is summed across all episodes for each individual state	N/A
Positive Urgency & Negative Urgency	PU & NU	Impulse responding to positive mood Impulse responding to negative mood	8	Need all items to calculate	N/A (PU and NU to be combined)	1-4	Sum	8-32
Behavioural Activation in Depression Scale- Short Form	BADS-SF	Behavioural avoidance vs approach	9	Need all items	Two	0-6	Sum Items 1,6,7,8 reverse coded	0-54
Behavioural Activation in Depression Scale- Short Form	BADS-SF Activation	Activation	5	Need all items	2, 3, 4, 5, 9	0-6	Sum	0-30
Behavioural Activation in Depression Scale- Short Form	BADS-SF Avoidance	Avoidance	4	Need all items	1,6,7,8	0-6	Sum Not reverse coded for subscale	0-24
Satisfaction with therapy[1]	N/A	Satisfaction with therapy	1	N/A	N/A	1-4	NA	1-4
Satisfaction with therapy	N/A	Acceptability of therapy	1	N/A	N/A	1-4	Sum	1-4

Questionnaire name	Acronym	Measured construct	N items	Maximum number of missing items for calculation of overall score	Subscales (N) Or Items in subscale	Individual item score (range)	Method of calculation of overall score	Overall score (range)
Satisfaction with therapy	N/A	Recommend to a friend	1	N/A	N/A	1-4	Sum	1-4
Satisfaction with therapy	N/A	Overall satisfaction with research	1	N/A	N/A	1-4	Sum	1-4
Adherence to medication	BARS[1]	Brief Adherence Rating Scale	1	N/A	N/A	0-100	NA	1-4
Ranking of clinical outcomes	N/A	Ranking of clinical outcome by participant priority: PHQ-9, ASRM, GAD, BRQ, QoLBD, ALS	N/A	Need to rank all six scales	N/A	1 to 6	Sum, rank	For each scale: take mean of score across all participants. Then rank mean scores from 1 to 6 1 is highest priority; 6 is lowest priority

ASRM: Altman Scale for Rating Mania. [1]Not applicable if participant is not using any medication.

Table A.2 Demographic data for individuals undergoing screening

Demographic characteristic	**All screened individuals (N=x)**
Age; mean (SD), median [min, max]	
Gender (categories TBC)	
Ethnicity	
White	
Mixed/multiple	
Asian/Asian British	
Black/African/Caribbean/Black British	
Other	
	Screened but not progressed to randomisation
Age; mean (SD), median [min, max]	
Gender (categories TBC)	
Ethnicity	
White	
Mixed/multiple	
Asian/Asian British	
Black/African/Caribbean/Black British	
Other	

[1]Characteristics are reported as number (percentage) unless otherwise stated.

Table A.3 Demographic data at baseline

Demographic characteristic[1]	**STABILISE+TAU (N=x)**	**TAU only (N=x)**
Age; mean (SD), median [min, max]		
Gender (categories TBC)		
Ethnicity		
White		
Mixed/multiple		
Asian/Asian British		
Black/African/Caribbean/Black British		
Other		
Perception of income		
Living comfortably on present income		
Coping on present income		
Finding it difficult on present income		
Finding it very difficult on present income		
Sexual Orientation		
Bisexual		
Gay man		
Gay woman/lesbian		
Heterosexual/straight		
Prefer not to say		
Prefer to self-describe		
Prescribed medication (categories TBC)		
Usual living situation		
Living alone (+/- children)		

Living with husband/wife (+/- children)		
Living together as a couple (+/- children)		
Living with parents		
Living with other relatives		
Living with others		
What kind of accommodation do you live in?		
Owner occupied flat or house		
Privately rented flat or house		
Flat or house rented from local authority/municipality or housing association/co-operative		
Nursing home		
Residential home		
Sheltered accommodation		
Other		
Currently open to secondary care services?		
Yes		
No		
Current employment status		
Paid employment		
Self-employment		
Unemployed		
Housewife/husband		
Student		
Retired		
Voluntary employment		
Sheltered employment		
Other		
Income		
Receive state benefits		
Yes		
No		

Benefits received		
Unemployment/income support		
Sickness/disability		
Housing		
State pension		
Child benefit		
Other		
Main income source		
Salary/wage		
State benefits		
Pension		
Family support		
Other		
Net income (£/month); mean (SD), n		

[1]Characteristics are reported as number (percentage) unless otherwise stated.

Table A.4 Feasibility outcomes

Outcome	Total cohort (N=x)	STABILISE+TAU (N=x)	TAU only (N=x)
Participant recruitment (participants per month); mean (SD), median [min, max]			
Devon Partnership NHS Trust		NA	NA
Avon and Wiltshire NHS Trust		NA	NA
Both sites		NA	NA
Retention at 14 weeks (any contact with trial team); n (%)			
Completion of first presented clinical outcome measure at 14 weeks; n (%)			
Completion of all four potential primary clinical outcome measures at 14 weeks; n (%)			
Retention at 30 weeks (any contact with trial team); n (%)			
Completion of first presented clinical outcome measure at 30 weeks; n (%)			
Completion of all four potential primary clinical outcome measures at 30 weeks; n (%)			
Retention at 52 weeks (any contact with trial team); n (%)			
Completion of first presented clinical outcome measure at 52 weeks; n (%)			
Completion of all four potential primary clinical outcome measures at 52 weeks; n (%)			
Full completion of momentary assessment data at 14 weeks; n (%)			

Percentage completion of momentary assessment at 14 weeks; mean (SD), median [min, max]			
Full completion of momentary assessment data at 30 weeks; n (%)			
Percentage completion of momentary assessment at 30 weeks; mean (SD), median [min, max]			
Attendance at a minimum of six STABILISE sessions	NA		NA

Table A.5 Clinical outcomes at baseline and follow-up

Outcome[1]	Baseline		14-week follow-up		30-week follow-up			52-week follow-up		
	STABILISE+TAU	TAU only	STABILISE +TAU	TAU only	STABILISE +TAU	TAU only	Between group mean difference (95% CI)	STABILISE +TAU	TAU only	Between group mean difference (95% CI)
PHQ-9										
GAD-7			NR	NR						
ALS Total										
ALS Anxiety Depression										
ALS Depression Elation										
ALS Anger										
BRMS			NR	NR						
QoLBD			NR	NR						
BRQ			NR	NR						
BARS								NR	NR	NR
Life Chart Self Report			NR	NR						
Mania			NR	NR						
N episodes			NR	NR						
N days			NR	NR						
Hypomania			NR	NR						
N episodes			NR	NR						
N days			NR	NR						
Depression			NR	NR						
N episodes			NR	NR						
N days			NR	NR						

[1]Reported as mean (SD), n. ALS: Affective Lability Scales, short version; BARS: Brief Adherence Rating Scale; BRMS: Bech–Rafaelson Mania Scale; BRQ: Bipolar Recovery Questionnaire; GAD-7: Generalised Anxiety Disorder 7-item; NR: Not reported; PHQ-9; Patient Health Questionnaire 9-item; QoLBD: Quality of Life Bipolar Disorder.

Table A.6 Process outcomes at baseline and follow-up

Outcome[1]	**Baseline**		**14-week follow-up**		**30-week follow-up**		
	STABILISE+TAU (n=x)	TAU only (n=x)	STABILISE+TAU (n=x)	TAU only (n=x)	STABILISE+ TAU (n=x)	TAU only (n=x)	Between group mean difference (95% CI)
PU&NU							
BADS-SF							
BADS-SF Activation							
BADS-SF Avoidance							

[1]Reported as mean (SD), n. BADS-SF: Behavioural Activation in Depression Scale- Short Form. PU&NU: Positive Urgency & Negative Urgency.

Table A.7 Satisfaction measures at 30-week follow-up

Satisfaction measure[1]	**STABILISE+ TAU (n=x)**	**TAU only (n=x)**	**Between group mean difference (95% CI)**
Satisfaction with therapy			
Acceptability of therapy			
Recommend to friend?			
Overall satisfaction with research			

[1]Reported as mean (SD), n.

Table A.8 Non-progression, withdrawal and loss to follow-up of potential participants and participants during trial procedures

Potential participants		
Stage Reason for non-progression		
Telephone screening	**Potential participants attending telephone screening (N=x)**	
Ineligible		
Potential participant choice		
Intake screening	**Potential participants attending intake screening (N=x)**	
Ineligible		
Potential participant choice		
Participants (N=x)		
Stage Change in participation status	**STABILISE+TAU (n=x)**	**TAU only (n=x)**
Randomisation to 14-week follow-up		
New withdrawals		
Reason 1 (etc)		
Passive non-contact[1]		
14-week follow-up to 30-week follow-up		
New withdrawals		
Reason 1 (etc)		
Cumulative total withdrawn		
Passive non-contact[1]		
30-week follow-up to 52-week follow-up		
New withdrawals		
Reason 1 (etc)		
Cumulative total withdrawn		
Passive non-contact[1]		

[1]Passive non-contact is timepoint specific.

Table A.9 Timing of outcomes data collection

Outcome	14-week follow-up		30-week follow-up		52-week follow-up	
	STABILISE+TAU (n=x)	TAU only (n=x)	STABILISE+TAU (n=x)	TAU only (n=x)	STABILISE+TAU (n=x)	TAU only (n=x)
Time to follow-up (days)[1]						
Completion of first potential primary clinical outcome measure within window[2]						
Completion of first potential primary clinical outcome measure outside window[2]						

[1]Reported as mean (SD), n; median [IQR; min, max]. Taken as time to completion of first potential primary clinical outcome measure at follow-up time. [2]Up to 4 weeks following scheduled date of collection.

Table A.10 Attendance at therapy sessions

Therapy sessions attended	**STABILISE + TAU (n=x)**
Initial sessions (maximum 22): Mean (SD); median [IQR]	
Consolidation sessions (maximum 3): Mean (SD); median [IQR]	
Total sessions (maximum 25): Mean (SD); median [IQR]	
Number of initial sessions attended; n (%)	
0	
1 to 5	
6 to 10	
11 to 15	
16 to 20	
21-22	
Number of consolidation sessions attended; n (%)	
0	
1	
2	
3	
Total number of sessions attended; n (%)	
0	
1 to 5	
6 to 10	
11 to 15	
16 to 22	
23-25 (at least one booster session)	

Table A.11 Intercurrent events

	STABILISE+TAU (n=x)	TAU only (n=x)
Receipt of therapy for BPD outside trial[1]		

[1]Reported as numbers and percentages.

Table A.12 Serious adverse events during 52-week follow-up

	At least one STABILISE session (n=x)	No STABILISE sessions (n=x)	Risk difference (95% CI)[2]	Odds ratio (95% CI)[2]
Death				
Self-harm incident (individual participants with at least one event)[1]				
Self-harm incident (number of events)			NA	NA
Hospital admission - physical health (individual participants with at least one event)[1]				
Hospital admission - physical health (number of events)			NA	NA
Hospital admission - mental health (individual participants with at least one event)[1]				
Hospital admission - mental health (number of events)			NA	NA
All adverse events (individual participants with at least one event)				
All adverse events (number of events)			NA	NA

[1]Reported as numbers and percentages. [2]At least one STABILISE session vs no STABILISE sessions.