TALKING WITH VOICES

NHS Greater Manchester Mental Health NHS Foundation Trust

# A novel dialogical therapy (Talking with Voices) in comparison to treatment as usual in adults with distressing and persistent auditory hallucinations: A randomised controlled trial to investigate the efficacy of a treatment strategy targeted at trauma-related mechanisms

Statistical Analysis Plan

Version 0.1 started: 13/07/2023

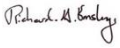
Based on v1.6 of the Talking with Voices protocol

ISRCTN 15897915

Junior Trial Statistician: Callum Glen

Signed: Date: 24/04/2024

Senior Trial Statistician: Richard Emsley

Signed:  Date: 08/05/25

Co-Chief Investigator: Eleanor Longden

Signed: Date: 28/01/2025

Co-Chief Investigator: Anthony Morrison

Signed: Date: 22/04/2025

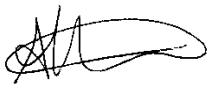


Table of Contents

# 1 Description of the trial

Talking with Voices aims to assess the clinical efficacy and mechanisms of a psychological therapy (Talking with Voices, TwV) for adults with Severe Mental Health Problems (SMHP) who hear persistent, distressing voices compared to treatment as usual, using the change in the Questionnaire about the Process of Recovery (QPR) scores as the primary outcome at 8-months post randomisation (post-treatment timepoint).

## 1.1 Principal research objectives to be addressed

### *1.1.1 Primary objective*

The primary objective is to compare the treatments (TwV + TAU vs TAU alone) on the change in total QPR score from baseline to 8 months.

### *1.1.2 Secondary Objectives*

Secondary objectives are to compare the treatments on both clinical efficacy and mechanistic properties, using the below measures:

- Psychotic Symptoms Rating Scale , PSYRATS-AH (Haddock et al. 1999)
- Voices Acceptance and Action Scale, VAAS (Brockman, Kiernan, and Murrell 2015)
- PSYRATS: Multimodal Hallucinations, (unpublished)
- PTSD Checklist for DSM-5, PCL-5 (Blevins et al. 2015)
- Brief Core Schema Scale: Self Subscale, BCSS (Fowler et al. 2006)
- Revised Dissociative Experiences Scale, DES-II (Carlson and Putnam 1993)
- The Revised Beliefs about Voices Questionnaire, BAVQ-R (Chadwick, Lees, and Birchwood 2000)
- Approve – Voices Questionnaire, (Hayward et al. 2020).

## 1.2 Trial Design (including blinding)

TwV is a 2 arm, rater-blinded, multisite, randomised controlled trial, recruiting participants that have heard voices for at least 1 year. We aim to recruit 296 participants.

The control group will receive treatment as usual, and the intervention group will receive the Talking with Voices intervention, alongside treatment as usual.

The primary endpoint is the end of treatment, at 8-months post treatment commencement for all recruited participants.

Outcome measures are collected at baseline, 8- and 14-months. It is expected that 207 participants will be followed up to 14 months. This means that the follow-up period will vary between 8-months to 14-months for participants, to maximise the recruitment period. These are estimated to be the first 207 participants recruited into the study, as collection of the 14-month follow-up is dependent on when participants are recruited into the trial. The aim of only

collecting the final data point for some participants is to maximise the recruitment period within a finite trial duration.

The trial is rater/assessor blinded. Assessment data will be collected by RA's who are blind to participants' allocation. The trial statistician will be blinded during the preparation of this document, and will then become unblinded to prepare the first closed DMC report, which will be approximately 6 months from the commencement of recruitment (in accordance with the KCTU SOPs, ST-06 v3,1). The senior statistician will be blinded throughout the trial. The investigators and research team will be blinded throughout the trial.

### *1.2.1 Figure 1- TwV Flow chart*

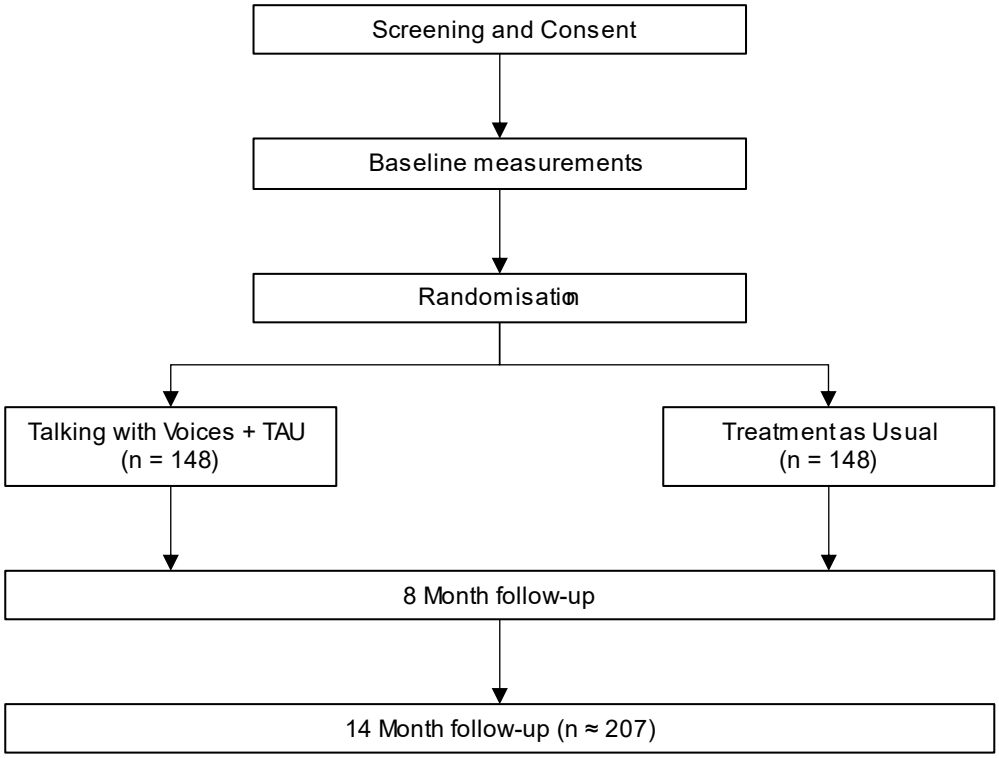



Figure 1 – Flow chart for participants through the TwV trial.

### 1.3 Method of allocation of groups

Randomisation will be in a 1:1 allocation to Treatment as Usual (TAU) or the Talking with Voices therapy plus TAU. Following informed and written consent, eligible participants will be randomised within 2 working days.

KTCU will hold and maintain a web-based randomisation system. Participants will be randomised at the individual level via randomisation lists generated using random permuted blocks, stratified by site (Greater Manchester, London, Newcastle and Oxford) and SMHP diagnosis (Schizophrenia spectrum diagnosis/early psychosis (F20 – F29), all other diagnosis) and administered via a study-specific web-based portal hosted by KCTU. Allocations will be made known by email to the Trial Manager (in order to monitor adherence to the randomisation algorithm), the Trial Administrator, trial therapists, site leads and CIs. The allocation will also be made known to participants by letter and phone call, and to relevant members of their healthcare team by letter. Blinding of the allocation code will be maintained for RAs until all outcome measures for all participants have been collected.

### 1.4 Duration of intervention period

The intervention period is 8-months.

Participants in the control arm will receive treatment as usual. TAU for SMHP is based on the Care Programme Approach, and can include psychiatric medication, assignment of community-based health and social care staff, care coordination, access to rehabilitative services, and outpatient care. With the exception of emergent risk issues, TAU alone will also not involve liaison between researchers and the participants' healthcare teams.

For participants in the intervention arm, they will receive the Talking with Voices therapy alongside treatment as usual. The treatment manual that was devised and refined during the TwV pilot (Longden et al. 2021), utilising individualised formulations ('constructs') to identify key psychosocial conflicts associated with voice hearing and determine targeted treatment strategies and shared goals for relational change. An 8-month treatment window permits up to 26 sessions, with the option of 4 additional booster sessions to consolidate therapeutic gains. Adherence checklists and electronic session records will be utilised to maximise fidelity to the manual, with any protocol divergences monitored during therapist supervision. Important treatment milestones will also be assessed and monitored.

### 1.5 Frequency and duration of follow-up and visit windows

There are 2 follow-up visits in the trial. The first is at 8-months post randomisation, which is the primary endpoint of the trial. The second is at 14-months post randomisation, a secondary endpoint. Not all participants in the trial will be followed up to 14-months. Whether a participant is followed up to 14-months or not will depend on when the participant was recruited into the trial. Those recruited towards the end of the trial will not be followed up to the 14-month timepoint. It is expected that approximately the first 207 participants will be followed up to 14-months. This number may change based on the rate of recruitment.

The date of randomisation is defined as day 1 for each participant in the trial.

The visit windows for each of the follow-up points are the date of the expected visit +60 days which means that the windows for visits are:

- 8-months, (243 days post randomisation, window of 243 – 303 days post randomisation)
- 14-months (425 days post randomisation, window of 425 – 485 days post randomisation)

This means that the earliest date that a participant can complete their follow-up assessment is the date which their follow-up is scheduled for. They can complete their follow-up visit up to 60 days (around 2-months) after their scheduled visit date, and the data will be still valid. If conducted outside of these timeframes, then data will be invalid.

## 1.6 Data collection

Outcome measures are collected from participants at baseline and at the follow-up visits. This collection will be carried out by Research Assistants (RAs) who are independent and blinded to the allocation of treatments. Collection will be done using a self-reported questionnaire for the primary outcome, and then a combination of self-reported questionnaires and structured interviews for secondary outcomes and mediation variables. The RAs will be trained as assessors and arrangements will be in place to ensure inter-rater reliability across the sites in the trial.

### *1.6.1 Eligibility criteria*

To be eligible for the trial, participants must fulfil all the inclusion criteria, and none of the exclusion criteria. If a potential participant does not fulfil one or more of the inclusion criteria, or does fulfil one or more of the exclusion criteria, then they are ineligible for the trial and therefore cannot be consented and randomised.

#### *1.6.1.1 Inclusion criteria*

1. Aged ≥ 16 years.
2. Heard voices for at least a year.
3. Scoring ≥ 1 on item 8 of the Psychotic Symptom Rating Scales – Auditory Hallucinations Subscale (PSYRATS-AH) (Haddock et al. 1999).
4. Able to provide written informed consent.
5. Actively help-seeking in relation to distressing voices.
6. In contact with mental health services for ≥ 6 months.
7. Willing and able to communicate with their voices and relay what the voices say to a therapist.
8. Hear voices that are sufficiently personified to engage in dialogical work.

#### *1.6.1.2 Exclusion criteria*

1. At immediate risk of harm to self or others.
2. Currently receiving structured, individualised psychological therapy.
3. Non-English speaking.
4. Primary diagnosis of alcohol/substance dependence or autism spectrum disorder.

5. Moderate/severe learning disability.
6. Organic cause for Voice Hearing.
7. Homeless/of no fixed abode.

### *1.6.2 Primary outcome measure*

The primary endpoint of the trial is 8-months post-randomisation, which is the end of treatment in both arms. The primary outcome is the total score on the 15-item Questionnaire about the Process of Recovery (QPR), which will be summarised as the group mean at each timepoint.

### *1.6.3 Secondary outcome measures*

There are multiple secondary outcome measures for the trial. These are listed below:

- PSYRATS-AH
- VAAS-12
- PSYRATS: Multimodal Hallucinations (Distress and Frequency continuous scales)
- PCL-5
- TVAQ (to be displayed descriptively at each timepoint)

These outcomes are all continuous scales, unless otherwise stated.

The mechanistic outcomes are:

- BCSS: Self Subscale
- DES-II
- BAVQ-R
- Approve – Voices Questionnaire

These outcomes are all continuous scales.

### *1.6.4 Baseline and demographic variables*

All outcomes will be measured at baseline. Additionally, the Trauma and Life Events checklist, TALE (Carr, Hardy, and Fornells-Ambrojo 2018) will be administered at baseline alone to assess the trauma participants may have faced.

The following demographic information will be collected from participants at baseline:

- Age at date of consent (Years)
- Gender
- Highest level of education,
- Main employment status (F/T, P/T, student, etc.),
- Marital status,
- Living arrangements,
- Ethnicity,
- Religion/belief,

- How long the participants have heard voices,
- Current psychiatric diagnosis,
- Any other psychiatric medications being taken,
- Any forms of psychological therapy partaken for voice hearing.

#### *1.6.5 Adverse events*

Adverse Events (AEs) and Serious Adverse Events (SAEs) will be collected for all trial participants from the time of their enrolment into the trial. The time of enrolment is defined as the date that the participant signs the consent form. Researchers at each site will note the AEs/SAEs at each follow up and enter into the MACRO database. The AEs will be logged in the unblinded database for TwV, therefore an unblinded member of the team will need to enter these events. If at the end of their follow up period, a participant has not experienced any AEs or SAEs, then the unblinded researcher will go into the database and enter that they have not experienced any AEs/SAEs in the duration of the trial. Any SAEs (or suspected SAEs) that are recorded will be reported to the Trial Coordinator.

All SAEs must be reported to the MHRA by the sponsor or a delegated to an appropriately qualified member of the research team. The relatedness of the SAE to the trial procedures and the trial intervention will be determined by the unblinded members of the research team. All SAEs that are determined to be related to the trial procedure or intervention and all SAEs that are determined to be unexpected will be reported to the REC.

All AEs/SAEs will be summarised and reported in the Open report to the DMEC. SAEs will also be shown by arm (partially blind, A/B, unless the DMEC members request to be fully unblinded) to the DMEC in the closed report. Action will then be taken accordingly depending on implication for the conduct of the trial.

The information to be recorded on AEs/SAEs will be:

- The system organ class,
- If a psychiatric disorder, the type of adverse event within that body system code,
- A description of the event,
- The start date of the event,
- The status of the adverse event (ongoing, recovered, death),
- The stop date of the event,
- The severity of the event,
- The relationship of the event to the trial procedures,
- The relationship of the event to the trial intervention,
- The serious or non-serious nature of the event.

### 1.7 Sample size estimation

The trial is a partially nested design, as there is clustering in the intervention arm due to the therapists giving the TwV therapy. Participants in the control arm therefore are considered as cluster sizes of 1.

The sample size calculation allows for 14 therapists over the course of the trial, with an Intraclass Correlation Coefficient (ICC) of 0.02, and each therapist to see an average of 9

participants, and variation in the cluster size of 9. This assumes that cluster membership follows a Poisson process.

Aiming to achieve 90% power to detect a between-group standard effect size (SES) of 0.4 at 8 months on the primary outcome (QPR), with the type-1 error set at two sided 5%, assuming a conservative correlation of 0.4 between the respective baseline and 8-month scores and a 1:1 allocation ration, 252 participants with outcome data would be required.

Accounting for a conservative 15% attrition of participants (the pilot trial saw 10% attrition rate) would then require 296 participants to be recruited. A recent study used an anchor-based method to establish the minimum important difference for the QPR and suggested that a difference of 4 – 5 points is a worthwhile target difference.

Using a difference of 4.5 points, with a standard deviation of 11.5 (based on QPR scores from several SMHP trials) this equates to an SES of 0.4, as above. In the TwV pilot trial, the observed SES was 0.7.

## 2.0 Data analysis plan – Data description

### 2.1 Recruitment

To show the flow of participants through the trial, a CONSORT diagram will be constructed (figure 2).

The CONSORT diagram includes the number of eligible patients, the number of patients that agree to enter the trial and the number that refuse to enter the trial. By treatment arm, the number that attend each follow-up visit, and those lost to follow-up/excluded/withdrawn at each follow-up visit. The numbers that adhered to the interventions will also be displayed.

Adherence to the intervention, TwV, is defined as receiving 8 or more sessions of the intervention.

### *2.1.1 Figure 2 – Template CONSORT diagram*

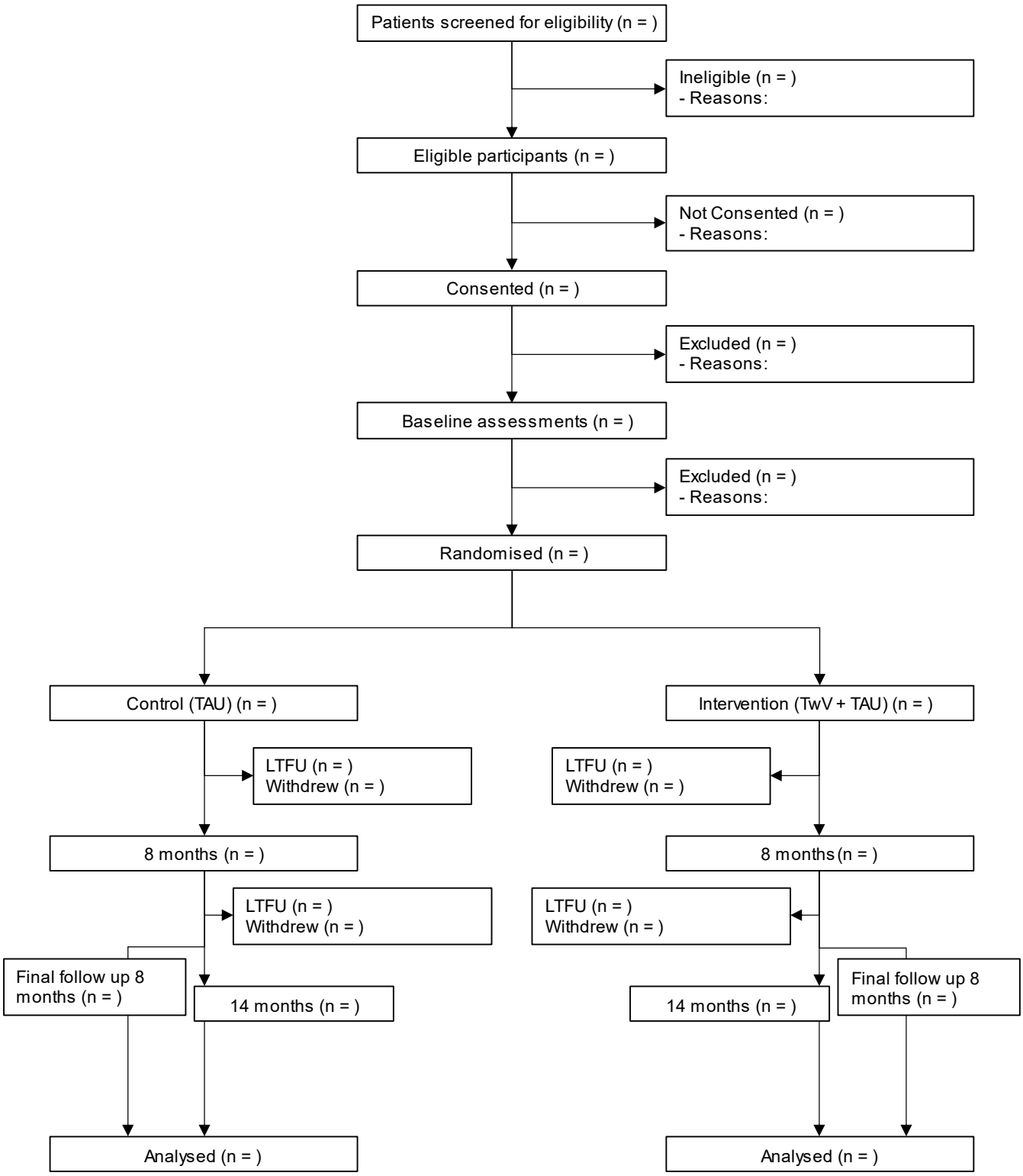



Figure 2 – Blank CONSORT diagram for the TwV trial.

## 2.2 Baseline comparability of randomised groups

Baseline descriptions of participants by arm and overall will be presented as means and standard deviation, median and inter-quartile ranges or number and proportion, as appropriate.

The descriptions will be of the demographic variables, as well as the outcome scores at baseline.

There will be no significance testing for baseline differences between groups. This is due to any differences being by chance, as is expected from the randomisation procedure.

## 2.3 Adherence to allocated treatment

Adherence to the intervention (TwV) is defined as attending 8 or more sessions of therapy. This is the same as was set in the pilot trial, where 87.5% of participants were adherent to the therapy.

Reasons for losses to follow up or withdrawals will be summarised, where available.

## 2.4 Loss to follow-up and other missing data

The number of participants who have missing data for each of the outcomes, primary or secondary will be summarised by arm and by timepoint.

The number of participants who withdraw from the trial will be summarised in a similar way, by arm and by timepoint.

It will be checked if compliance with the intervention is predictive of missingness at later timepoints, as this could inform the missing data approach that is used.

The reasons for withdrawals and losses to follow-up will be summarised (if known).

## 2.5 Adverse event reporting

Adverse Events (AEs) and Serious Adverse Events (SAEs) will be summarised by arm. The relatedness of the AEs and the SAEs to the trial intervention will also be summarised, by arm.

There will be a line listing of the serious adverse events, by arm.

## 2.6 Assessment of outcome measures (unblinding)

The outcome assessors will be blind to allocation. The senior statistician will remain blinded throughout the trial. The trial statistician will be blinded until the signoff of this document is completed.

## 2.7 Descriptive statistics for outcome measures

The primary and secondary outcome measures will be summarised using the appropriate summary statistics, overall and by arm, at each timepoint that they are collected.

The distributions of the continuous outcome measures will be inspected, where they will be checked that they are normally distributed.

The unadjusted means of the outcomes will be plotted over time for each measure.

# 3.0 Data analysis plan

## 3.1 Primary analysis description

### *3.1.1 Definition of primary estimand*

The primary estimand used will be a treatment policy estimand. This means that the occurrence of any Intercurrent Event (IE) will be not be accounted for in the primary analysis. The treatment policy estimand is very similar to an intention to treat analysis.

IEs which could potentially occur in this trial are:

- Death,
- Non-attendance at therapy,
- Receiving another therapy outside of the trial,
- Hospitalisation of participants for any reason,
- Hospitalisation of participants for mental health reasons,
- Diagnosis of another SMHP,
- Circumstances where participants are unable to attend therapy for long periods,

Using the treatment policy estimand, if a participant experiences any one or more of these IEs during their involvement in the trial, all data that they provide will still be used in the final analysis of the trial.

The estimand as a whole is defined using the following attributes:

- **Population**: participants with SMHPs who have heard voices for at least a year that met the inclusion criteria for the trial and none of the exclusion criteria.
- **Endpoint**: Primary endpoint is 8-months post randomisation, at the end of treatment.

- **Treatment condition**: TwV and TAU will be compared against TAU alone.
- **Intercurrent events**: As listed above
- **Population level summary**: between group difference in mean QPR score at 8-months in those receiving TwV and TAU against those receiving TAU alone, with adjustment for baseline scores.

We will report the incidence of IEs by randomised group.

### *3.1.2 Treatment policy analysis set*

The primary analysis will include all randomised participants who are eligible, in the arm of the trial which they are randomsied to, regardless of the treatment which they actually received during the trial. This approach using a treatment policy estimand is very similar to the traditional intention to treat (ITT) analysis. This estimand will be used for the primary, and all secondary outcomes, unless otherwise stated.

### *3.1.3 Analysis of primary outcome*

A linear mixed model will be fit to the data to compare those in the TwV + TAU arm against the TAU arm alone.

The model will include fixed effects for treatment, site, time (categorical), baseline measure of the outcome, diagnosis (Schizophrenia spectrum diagnoses/early psychosis (F20 – F29) or all other diagnoses) and a treatment by time interaction.

The clustering effect of the therapist in the TwV arm of the trial will be included in the model using a random intercept for therapist for those in the TwV arm only. A random intercept will be included in the model to account for the clustering of observations at each timepoint within participants.

The results presented will be the difference in the QPR score at each follow-up point as the estimate of treatment effect, alongside the 95% confidence intervals for the estimates. Estimates will be provided adjusted and unadjusted. If the treatment effect estimate on the primary outcome (QPR) is significant at the 5% ($\alpha = 0.05$), two-sided level in the analysis, then the intervention will be deemed superior to the control.

### *3.1.4 Analysis of secondary outcomes*

Analysis of the secondary outcomes will be carried out using similar models as for the primary outcome.

For those secondary outcomes which are only collected at one timepoint, then there will be no random intercept for participant as there will be no clustering of observations within a participant.

## 3.2 Statistical considerations

### *3.2.1 Time points*

Data is admissible into the analysis if the outcome was collected within the visit window specified for each of the visits (not admissible before visit date, admissible up to 60 days after the scheduled visit).

### *3.2.2 Stratification and clustering*

The randomisation procedure in this trial is stratified by both site and SMHP diagnosis. Therefore, these two factors are included in the analysis model.

The structure of the data is longitudinal, with repeated measures (baseline, 8- and 14-months). The correlation of observations within participants will be accounted for in the analysis model by including a random intercept for participant.

The participants in the TwV arm are also clustered by the therapist who delivered their TwV therapy. This is accounted for in the analysis models by including a random intercept for therapist in the TwV arm only, which allows for clustering and potential correlation between participants by therapist, and as well as the clustering and potential correlation between observations within each participant.

### *3.2.3 Missing items in scales and subscales*

The number and percentage of participants with complete data will be reported.

If there is missing items guidance for a measure/scale, then this guidance will be used to deal with the missing items.

If there is no guidance for missing items, then scales will be pro-rated for a participant if 20% or fewer of the items in the scale are missing. This works by taking, for example, a scale of 10 items, where participants can have up to 2 items missing. If a participant has 2 or less items missing, then the mean of the items that were answered is calculated and used to replace the missing item values. The score for that participant is then calculated based on their already completed values and the replacement values for missing items.

### *3.2.4 Missing baseline data*

Missing baseline data should not be an issue for this trial. Baseline data is collected upon consent into the trial, so all participants should have complete baseline data. If there is any missing baseline data for the baseline scores, then the outcomes will be scored as defined in section 3.2.3 if possible, or using simple mean imputation if this is not possible..

Some analyses may use other baseline variables and 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 (White and Thompson 2005).

### *3.2.5 Missing outcome data*

Missing outcome data will be dealt with by fitting linear mixed models to all available data using maximum likelihood methods. Using this approach provides valid inferences under the assumption that the missing data mechanism is ignorable, or that data is missing at random (MAR). If post treatment variables such as compliance with treatment are found to be predictive of drop-out, then multiple imputation will be considered.

### *3.2.6 Method for handling multiple comparisons*

The trial has one prespecified primary outcome and analysis. The secondary outcomes and their analyses can be considered exploratory, therefore it is not necessary to adjust for multiple comparisons.

### *3.2.7 Methods for handling non-compliance*

In addition to the primary analysis using the primary estimand, the effect of adhering to treatment for the TwV arm will be estimated via a supplementary analysis.

Adherence for the TwV therapy is defined as attending 8 or more (≥ 8) sessions of the therapy.

A complier average causal effect (CACE) model will be used to deal with adherence in the population. The population for the CACE analysis is those participants, irrespective of their allocation, would always have complied with the intervention had they been given the chance to.

### *3.2.8 Model assumption checks*

Common assumptions for mixed models will be checked for the primary outcome.

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, the models will be fit using robust standard errors.

Any model assumption issues that lead to a change in the planned analysis will be reviewed by the blinded senior statistician and approved prior to implementation.

### *3.2.9 Sensitivity analyses*

Discuss with trial team, none specified in the protocol.

Potential sensitivity analyses:

MNAR sensitivity analysis: It may be possible that missing data is missing not at random (MNAR). For the primary outcome, we will use a range of delta values (i.e. on average, dropouts have a difference of X points on the QPR compared to those retained) and examine the estimated treatment effects under different values of X. An appropriate method may be utilised (e.x. mean-score method, or utilising delta after multiple imputation).

### *3.2.10 Planned subgroup analyses*

The subgroup analyses to be conducted will conducted as to compare schizophrenia spectrum diagnoses with the other diagnoses in the trial, to assess the primary outcome trans-diagnostically.

Other subgroups investigated will include gender, age and ethnicity. Subgroup analyses methods will match the primary analysis method.

### *3.2.11 Exploratory analysis*

No exploratory analysis is planned at this stage.

### *3.2.12 Mediation and Moderation*

After the primary analysis is conducted, if an effect is found, mediation analysis will be carried out.

The total effects will use the QPR at 8-months with independent variables, excluding the mediation variable.

The mediation models will then be fit, and the following will be reported from each of these models:

- a paths,
- the effect of intervention arm on the mediator as mean differences between the arms,
- b paths, as the change in QPR for a one point change in the mediator,
- c' paths, as the mean difference in QPR between arms taking the mediator into account.

The indirect effect will be calculated by multiplying the a and b paths together. For the indirect effect, the bias-corrected and accelerated (bca)bootstrap CI from bootstrapping with 1000 repetitions will be used. This will be presented as a proportion of the total effect that mediated for each model (calculated as indirect effect estimate divided by the total effect estimate, multiplied by 100).

The potential mediators to be considered are the mechanistic secondary outcomes. These outcomes will not be considered as mediators if there is not an effect when they are tested as secondary outcomes.

A secondary mediation is also considered, using PSYRATS-AH as the outcome variable of interest. This relies on the secondary analysis of PSYRATS-AH finding an effect.

Moderation analyses are planned for socio-demographic measures, such as age, sex and gender.

## 3.3 Internal Pilot

In October 2024 (13 months of recruitment, 5 months of follow up), the progress of the trial will be assessed against pre-specified criteria, in the table below.

| **Criteria** | **Red** % (n) | **Amber** % (n) | **Green** % (n) |
|---|---|---|---|
| Trial recruitment | ≤59 (≤116) | 60-99 (116-192) | 100 (193) |
| Recruitment rate per month | ≤59 (≤8) | 60-99 (9-14) | 100 (15) |
| Number of sites opened | ≤50 (≤2) | 70 (3) | 100 (4) |
| Proportion receiving allocated intervention | ≤59 | 60-99 | 100 |
| Proportion with complete primary outcome data | ≤84 | 85-99 | 100 |

## 3.4 Software

### *3.3.1 Data management*

The database and data collection will use InferMed MACRO, held at KCL on a dedicated, secure server and managed by the KCTU. The KCTU data manager will extract data periodically as needed and provide these in the requested format (Stata or csv) to the analyst(s). No other members of the trial team should be requesting data extracts, particularly extracts that contain outcome data.

### *3.3.2 Statistical software*

Stata 18 (or a similar package) will be used to conduct all of the statistical analysis required throughout the duration of the trial.

## 4.0 Schedule of assessments and measures

| Assessment/Measure | Screening | Baseline | Month 8 | Month 14 |
|---|---|---|---|---|
| Eligibility | X | | | |
| Demographic form | | X | | |
| Modified Trauma and Life Events Checklist (TALE) | | X | | |
| The Questionnaire about Process of Recovery (QPR) | | X | X | X |
| The Psychotic Symptom Rating Scale: Auditory Hallucinations Subscale (PSYRATS-AH) | | X | X | X |
| Voices Acceptance and Action Scale (VAAS-12) | | X | X | X |
| Modified PSYRATS: Multimodal Hallucinations | | X | X | X |
| PTSD Checklist for the DSM-5 (PCL-5) | | X | X | X |
| Trauma Voice Associations Questionnaire (TVAQ) | | X | X | X |
| Brief Core Schema Scale (BCSS) Myself subscale | | X | X | X |
| Revised Dissociative Experiences Scale (DES-II) | | X | X | X |
| Beliefs About Voices Questionnaire – Revised (BAVQ-R) | | X | X | X |
| Approve – Voices Questionnaire | | X | X | X |

## 5.0 Statistical analysis plan amendments

The initial signed version of the SAP will be numbered version 1.0. Minor amendments to this SAP will be numbered v1.1, 1.2 etc. Major amendments will be numbered as v2.0, 3.0 etc.

After the initial SAP has been signed, if any changes are required, these will be decided upon by the blinded senior statistician. This is to avoid any chance of bias that could be introduced from the trial statistician, if they have seen unblinded data.

For all changes, a copy of the tracked changes version will be retained, and the updated version will be approved by the relevant parties that approved the initial version (CI, senior statistician, TSC statistician). Changes will be tracked in the below table.

| Version | Date of approval | Summary of changes |
|---|---|---|
| 1.0 | | Initial version |
| | | |
| | | |

# 6.0 Tables

## 6.1 Baseline summary of characteristics

Table 1 – Baseline characteristics of participants, by arm.

| Characteristic | Control | Trial arm Intervention | Total |
|---|---|---|---|
| N | XXX | XXX | XXX |
| Age | | | |
| Mean (SD) | XX.XX (XX.XX) | XX.XX (XX.XX) | XX.XX (XX.XX) |
| Median (IQR) | XX.XX (XX.XX) | XX.XX (XX.XX) | XX.XX (XX.XX) |
| Gender | | | |
| Male | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Female | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Transgender | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Non-binary | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Other | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Prefer not to answer | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Highest level of education | | | |
| Never attended school | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Primary | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Secondary | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Further | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Higher | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Main employment status | | | |
| Full time | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Part time | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Voluntary | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Unemployed | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Retired | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Student | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Marital status | | | |
| Single | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Married | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Divorced | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Separated | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Widowed | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Cohabitating | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Civil partnership | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Living arrangements | | | |
| Parent(s) only | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Parent(s) and siblings | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Partner | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Alone | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Friends | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Flatmates/housemates (not friends) | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Extended or other family | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Carer | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Other | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Ethnicity | | | |
| Bangladeshi | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |

| | | | |
|---|---|---|---|
| Indian | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Pakistani | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Chinese | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Any other Asian background | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| African | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Caribbean | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Any other Black background | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| White & Asian | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| White & Black African | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| White & Black Caribbean | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Any other mixed background | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| White British | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| White Irish | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Any other White background | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Any other ethnic group | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Prefer not to answer | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Religion/Belief | | | |
| Atheism (no belief) | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Buddhism | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Christianity | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Islam | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Jainism | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sikhism | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Judaism | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Hinduism | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Rastafarianism | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Other | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Length of voice hearing | | | |
| Mean (SD) | XX.XX (XX.XX) | XX.XX (XX.XX) | XX.XX (XX.XX) |
| Median (IQR) | XX.XX (XX.XX) | XX.XX (XX.XX) | XX.XX (XX.XX) |
| Current psychiatric diagnosis | | | |
| Schizophrenia | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Psychosis | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Schizoaffective disorder | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| PTSD | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Depression | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Bipolar disorder | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| EUPD | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Other | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Currently taking antipsychotic medication | | | |
| Yes | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| No | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Taking other psychiatric medication | | | |
| Yes | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| No | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Previously received psychological therapy for voice hearing | | | |
| Yes | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| No | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |

## 6.2 Baseline summary of scores

Table 2 – Participants baseline scores, by arm and overall.

| Score | Trial Arm | | |
|---|---|---|---|
| | Control | Intervention | Total |
| N | XXX | XXX | XXX |
| QPR | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| PSYRATS-AH * | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| VAAS-12 | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Acceptance subscale | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Action subscale | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| PSYRATS-MMH | | | |
| Occurrence of.. | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Distress caused... | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Bothered the most by... | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Combination | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Frequency of.... | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| PCL-5 | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| TVAQ | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Part A | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |

| | | | |
|---|---|---|---|
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Part B | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Part C Q-11 | | | |
| No | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Yes | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Part C Q-12 | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Part C Q-13 | | | |
| No | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Yes | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Part C Q-16 | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| BCSS | | | |
| Positive | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Negative | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| DES-II | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| BAVQ-R | | | |
| Malevolence | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Benevolence | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Omnipotence | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Resistance | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Engagement | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Approve – Voices Questionnaire | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Assertive relating | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Aggressive relating | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Passive relating | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |

* The PSYRATS will be described by individual items in the final report

### 6.3 8-month summary of scores

Table 3 – Summary of scores and measures at the 8-month follow up timepoint.

| Score | Trial Arm | | |
|---|---|---|---|
| | Control | Intervention | Total |
| N | XXX | XXX | XXX |
| QPR | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| PSYRATS-AH | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| VAAS-12 | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Multimodal Hallucinations | | | |
| Occurrence of.. | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Distress caused... | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Bothered the most by... | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Combination | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| PCL-5 | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| TVAQ | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| BCSS | | | |
| Positive | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Negative | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| DES-II | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| BAVQ-R | | | |

| | | | |
|---|---|---|---|
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Approve – Voices | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |

### 6.4 14-month summary of scores

Table 4 – Summary of scores and measures at the 14-month follow-up point.

| Score | Trial Arm | | |
|---|---|---|---|
| | Control | Intervention | Total |
| N | XXX | XXX | XXX |
| QPR | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| PSYRATS-AH | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| VAAS-12 | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Multimodal Hallucinations | | | |
| Occurrence of.. | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Distress caused... | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Bothered the most by... | | | |
| Auditory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Tactile | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Visual | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Olfactory | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Sensed presence | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| Combination | XXX (XX.XX%) | XXX (XX.XX%) | XXX (XX.XX%) |
| PCL-5 | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| TVAQ | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| BCSS | | | |
| Positive | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Negative | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| DES-II | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| BAVQ-R | | | |

| | | | |
|---|---|---|---|
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |
| Approve – Voices | | | |
| Mean (SD) | XX (XX) | XX (XX) | XX (XX) |
| Median (IQR) | XX (XX – XX) | XX (XX – XX) | XX (XX – XX) |

## 6.5 Primary outcome results

Table 5 – Results of the primary outcome (QPR) analysis, using the control group as the reference.

| Analysis | Estimated difference between arms | SE | 95% CI | P value |
|---|---|---|---|---|
| 8-month change | | | | |
| Unadjusted | X.XX | X.XX | X.XX – X.XX | 0.XXX |
| Adjusted | X.XX | X.XX | X.XX – X.XX | 0.XXX |
| 14-month change | | | | |
| Unadjusted | X.XX | X.XX | X.XX – X.XX | 0.XXX |
| Adjusted | X.XX | X.XX | X.XX – X.XX | 0.XXX |

A table similar to that seen above (table 5) will be replicated for the secondary outcome analyses.

# References


Blevins, Christy A, Frank W Weathers, Margaret T Davis, Tracy K Witte, and Jessica L Domino. 2015. 'The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation', *Journal of traumatic stress*, 28: 489-98.

Brockman, Robert, Michael Kiernan, and Elizabeth Murrell. 2015. 'Psychometric properties of two brief versions of the Voices Acceptance and Action Scale (VAAS): Implications for the second-wave and third-wave behavioural and cognitive approaches to auditory hallucinations', *Clinical psychology & psychotherapy*, 22: 450-59.

Carlson, Eve Bernstein, and Frank W Putnam. 1993. 'An update on the dissociative experiences scale', *Dissociation: progress in the dissociative disorders*.

Carr, Sarah, Amy Hardy, and Miriam Fornells-Ambrojo. 2018. 'The Trauma and Life Events (TALE) checklist: development of a tool for improving routine screening in people with psychosis', *European journal of psychotraumatology*, 9: 1512265.

Chadwick, Paul, Susan Lees, and MAX Birchwood. 2000. 'The revised beliefs about voices questionnaire (BAVQ–R)', *The British Journal of Psychiatry*, 177: 229-32.

Fowler, David, Daniel Freeman, BEN Smith, Elizabeth Kuipers, Paul Bebbington, Hannah Bashforth, Sian Coker, Joanne Hodgekins, Alison Gracie, and Graham Dunn. 2006. 'The Brief Core Schema Scales (BCSS): psychometric properties and associations with paranoia and grandiosity in non-clinical and psychosis samples', *Psychological medicine*, 36: 749-59.

Haddock, Gillian, J McCarron, N Tarrier, and EB Faragher. 1999. 'Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS)', *Psychological medicine*, 29: 879-89.

Hayward, Mark, Björn Schlier, Clara Strauss, Aikaterini Rammou, and Tania Lincoln. 2020. 'Construction and validation of the approve questionnaires–measures of relating to voices and other people', *Schizophrenia Research*, 220: 254-60.

Longden, Eleanor, Dirk Corstens, Anthony P Morrison, Amanda Larkin, Elizabeth Murphy, Natasha Holden, Ann Steele, Alison Branitsky, and Samantha Bowe. 2021. 'A treatment protocol to guide the delivery of dialogical engagement with auditory hallucinations: Experience from the talking with voices pilot trial', *Psychology and Psychotherapy: Theory, Research and Practice*, 94: 558-72.

White, Ian. R., and Simon. G. Thompson. 2005. 'Adjusting for partially missing baseline measurements in randomized trials', *Stat Med*, 24: 993-1007.

# Appendix

## Scale scoring procedures

### QPR

The Quality of Processes of Recovery is a 15 item scale, where the scores for each items are 0-4, 0 being not at all and 4 being extremely. The total score is calculated by summing the 15 items, with a score range of 0 to 60.

If participants have 3 or less items missing, their score will be pro-rated using the items they have completed. If there are more than 3 items missing, a score will not be calculated.

### PSYRATS – Auditory Hallucinations

The PSYRATS – AH is an 11 item scale, where each item is scored 0 (absent) to 4 (severe). The total score is calculated by summing the 11 items, with a score range from 0 to 44.

If participants have 2 or less items missing, their score will be pro-rated using the items they have completed. If there are more than 2 items missing, a score will not be calculated.

### PSYRATS – Multimodal Hallucinations

This measure will only be taken if the participant identifies an unusual experiences other than Auditory/Hearing. These can be,

- Tactile/somatic/Feeling
- Visual/seeing things
- Olfactory/gustatory/smelling
- Felt presence/felt sense

Each of the hallucinations can be given a frequency score (scored 0-4, 0 Not at all to 4 Constantly) and a distress score (scored 0-4, 0 Not at all to 4 Very much). Participants are also asked which of the experiences distresses them the most.

The frequency and distress scores will be summarised for each experience, for those who have experienced it, as no score will be calculated.

### VAAS

The Voices Acceptance and Action scale is a 12 item scale, where a 1-5 Likert scale is used for each question (1 = Strongly disagree, 5 = Strongly agree). The score is a sum of the 12 items, where items 4, 6, 7, 8 and 10 are reverse coded. Scores range from 12 – 60.

If more than 2 items are missing in a participants scale, their score will not be calculated. If there are 2 or less items missing, then their score will be pro-rated using the items they have completed.

### PCL-5

The PCL-5 is a 20-item scale where each item is scored 0-4. The total score is calculated by summing all 20 items, to give a total score with a range of 0 – 80.

If 4 or less items are missing, participants score will be pro-rated using the items that they have completed. If there are more than 4 items missing, then participants will not have a score calculated.

**TVAQ**

There is no way to score this outcome, therefore results will be displayed descriptively.

**DES-II**

The Dissociative Experiences Scale-II is a 28 item scale, where each item is scored 0-100. The items are summed, then the scores are averaged, to give a total score between 0 and 100.

If a participant has 5 or less items missing, then a score will be pro-rated using the items that they have completed. If there are more than 5 items missing, then a score will not be calculated.

**BCSS**

12 items, split into a 6 item negative scale and a 6 item positive scale. There is no total scale. Items are scored by summing together all the items in each subscale. Items are scored 1-4, the range of each subscale is 6-24.

If a participant has 1 item or less missing in each subscale respectively, then a score will be pro-rated using the items that they have completed. If there are more than 1 item missing, then a score will not be calculated for the subscale.

**BAVQ-R**

The BAVQ-R consists of 35 items, split into the following 5 subscales:

- Malevolence – 6 items, 6-24 score range
- Benevolence – 6 items, 6-24 score range
- Omnipotence – 6 items, 6-24 score range
- Resistance – 9 items, 9-36 score range
- Engagement – 8 items, 8-32 score range

Each subscale is pro-rated at 20%. If more than 20% of items for one subscale is missing, then a score will not be calculated. There is not total score for this outcome.

**Approve – Voices Questionnaire**

Approve Voices had 15 items, which are scored 0-10 each. The score is calculated by summing the items, to give a range of 0-150.

If less than 3 items are missing for a participant, then a score will be pro-rated. If there are more than 3 items missing, then there will not be a score calculated.