Approved: 11 October 2022
Statistical Analysis Plan

FULL TITLE	Managing Unusual Sensory Experiences (MUSE): A feasibility trial of a targeted, psycho-education toolkit for distressing hallucinations, in people with a first episode of psychosis (MUSE FEP)
REC Number:	21/YH/0090
SAP VERSION	3
ISRCTN NUMBER	16793301
SAP VERSION DATE	23/09/2022
TRIAL STATISTICIAN	James Liley
SENIOR STATISTICIAN	Emmanuel Ogundimu
TRIAL CHIEF INVESTIGATOR	Robert Dudley
SAP AUTHOR	Ehsan Kharati, James Liley

1 SAP Signatures

I give my approval for the attached SAP entitled "Managing Unusual Sensory Experiences (MUSE): A feasibility trial of a targeted, psycho-education toolkit for distressing hallucinations, in people with a first episode of psychosis" dated XXX

Chief Investigator

Dr Robert Dudley

Signature: 

Date: 11th October 2022

Statistician

Dr James Liley

Signature: 

Date: 26th October 2022

Senior Statistician

Dr Emmanuel Ogundimu

Signature: 

Date: 11/10/2022

2 Table of Contents

3 Abbreviations and Definitions

AVH	Auditory verbal hallucination
EIP	Early Intervention in Psychosis
CNTW	Cumbria, Northumberland, Tyne and Wear
CBT	Cognitive behaviour therapy
CONSORT	Consolidated standards of reporting trials
CPN	Community psychiatric nurses
LEAP	Lived experience advisory panel
MUSE	Managing unusual sensory experiences
NICE	National Institute for Health & Care Excellence
NHS	National health service
PPO	Pseudo-primary outcome
RCT	Randomised control trial
SAP	Statistical analysis plan
TAU	Treatment as usual
TEWV	Tees, Esk and Wear Valley

4 Introduction

4.1 Preface

Psychosis refers to a number of conditions – such as schizophrenia. Schizophrenia is among the top ten disorders in burden, disability, and societal and health costs worldwide(Hjorthøj et al., 2017; Kennedy et al., 2014). Almost all patients with psychosis at some stage have hallucinations (hearing or seeing things that others do not). These hallucinations can lead to withdrawal, with many people spending much of their time alone(Kennedy et al., 2014). Physical activity levels are reduced by about two thirds (Kennedy et al., 2014) and over 90% of patients with schizophrenia are unemployed. People with psychosis have an increased risk of early mortality and high suicide rates (Hjorthøj et al., 2017). In 2012 the total annual cost to the public sector in England was estimated at over £7 billion (Andrew et al., 2012).

Hallucinations are a common feature of psychosis, causing significant distress and disability. The National Institute for Health & Care Excellence (NICE) recommends that all individuals with psychosis be offered Cognitive Behaviour Therapy for psychosis (CBTp) (NICE., 2014). However, in practice access is often limited owing to a lack of CBTp-trained staff. One solution is to develop shorter, targeted treatments that use CBTp-informed techniques to focus on specific symptoms. This has produced promising results for other symptoms of psychosis, such as delusions (i.e. unusual beliefs), but there has been less research on hallucinations. Our toolkit, called MUSE (Managing Unusual Sensory Experiences) explains why people have hallucinations and helps the person to develop and use coping strategies to reduce distress. MUSE uses psychoeducation about the currently known causal mechanisms of hallucination as means of exploring, with service users, why their specific experiences may be happening. This knowledge is then matched to specific, tailored interventions and coping strategies that enable the person to understand and manage their experiences differently and reduce their distress. This process relies on psychoeducation as its basis, but it is more fundamentally about helping a person to change their understanding, manage their experiences better and thus cope more effectively. Unlike CBTp, MUSE intervention focuses only on hallucinations, and as such the treatment is short (4-6 one hour weekly sessions). It can be used on tablets, laptops or desktop computers (ensuring its accessibility to NHS staff) and provides information about hallucinations in a user-friendly and engaging way, including use of audio, video, and animated content. Crucially, it is designed for use not just by trained CBT therapists, but also non-specialist staff like Community Psychiatric Nurses (CPNs).

4.2 Purpose of the analyses

The long-term aim of this research is to increase the number of psychosis patients who have access to an intervention to reduce the distress of hallucinations. Prior to the long-term aim, effectiveness and cost-effectiveness of the intervention will be established. As a step towards this the proposed study aims to establish if it is feasible to conduct a future clinical- and cost-effectiveness study.

4.3 Checklist

Appendix 3 contains a checklist for this SAP against recommendations from the CONSORT 2010 guidelines (Eldridge et al., 2016).

5 Study Objectives

The main objective of the proposed study is to establish if it is possible to undertake a larger, definitive trial in the future. The proposed study will i) assess the feasibility of CPN staff training and delivery of the MUSE toolkit; ii) asses the acceptability of the MUSE toolkit to patients and staff; and to iii) collect data to inform the effectiveness of a future definitive trial.

Three work streams map onto the objectives and will establish:

1: Feasibility of staff training and delivery of the MUSE toolkit

We will explore whether the planned training equips the CPNs with the skills and confidence to deliver MUSE. This will be assessed through quantitative and qualitative evaluation at the end of training, six months into the study, and at end of the study. We will also explore whether the supervision package is sufficient and useful to support CPNs in the delivery of this toolkit (using the same methods and time-points).

2: Acceptability of the MUSE toolkit to patients and staff

We will also explore the acceptability of the intervention to the participants engaging in the treatment and to CPNs delivering it. At the follow up meeting, participants will be asked to share their views. Staff will be asked in the latter stages of the study their views on the toolkit. We will also investigate whether the intervention needs further refinement, by asking staff and service users at these time points their views on the toolkit.

3: We will inform the development of future definitive trial by collecting the following data:

- Proportion of eligible individuals that clinicians are willing to refer (referral rate)
- Proportion of eligible individuals willing to participate (recruitment rate) and the proportion of participants who comply with their allocation (allocation compliance rate)
- Proportion of participants who drop-out of the study (attrition rate)
- Characteristics of trial participants
- Appropriateness and integrity of treatment protocols according to CPNs and patients
- Completion rate of measures

- Acceptability, relevance and validity of the measures to assess clinical effectiveness and safety in a subsequent definitive trial

- Appropriateness of quality of life measures, and service use data needed to undertake a future full health economic evaluation

- Access to CBTp and MUSE like interventions in other EIP services in England

We will also explore the acceptability of the trial methods and procedures to the participants engaging in the study and to CPNs involved in supporting it.

6 General Study Design and Plan

The proposed study is a two-arm feasibility RCT comparing MUSE and treatment as usual (TAU) (n=40) to TAU alone (n=40), recruiting across two sites, using 1:1 allocation and blind assessments at baseline, 2 month and 3 months follow ups. There will be two recruiting sites in the North East of England (Cumbria, Northumberland, Tyne and Wear; Tees, Esk and Wear Valley), both of which are Foundation NHS mental health Trusts and serve large populations (CNTW population 1.7 million and TEWV 1.7 million). MUSE therapy will involve a patient meeting with a CPN over several weekly sessions – usually in their own home – to understand the causes and consequences of hallucinations (4 core sessions with an option of 2 extra). The study will use a mixed methods approach. Quantitative information on recruitment rates, adherence, and completion of outcome assessments will be collected. Participants will receive assessments pre- and post-treatment (six to eight weeks), and at three to four months follow-up. The assessments will measure distress and disability caused by hallucinations, depression, quality of life, perceived recovery, therapeutic relationship and intervention quality (the latter two will be measured at end of treatment and follow-up). These measures will not be used to determine effectiveness but to help identify important parameters for a future trial (i.e. completion rates and selection of best outcome measures). We nominate a pseudo-primary outcome (PPO) and list several potential alternative primary outcomes for a definitive study. These may be chosen over the PPO on the basis of adherence, convenience, acceptability, similarity to the PPO, and estimated effect size. We will estimate the minimum sample size necessary in a definitive study in order to detect the estimated effect size for the PPO. Statistical analysis will primarily use linear mixed models in which treatments are fixed effects, and will evaluate treatments according to the size and confidence intervals of coefficients of these effects.

6.1 Study intervention

The intervention we will use is a novel treatment manual for hallucinations. The treatment is divided into the following Modules:

1. What are Voices? This module provides normalising information about the frequency of voices and the factors that tend to increase voice-hearing (for example substance misuse and sleep deprivation), along with testimonies from other voice-hearers.

2. How the Mind Works. This module outlines current understanding of key psychological processes such as threat detection, the importance of prediction (top-down processing) and how intrusive thoughts work.

3. Assessment. This module identifies the subtype of hallucination a service-user is experiencing. After the assessment the therapist should be able to identify whether the voice-hearing is an Inner Speech-Auditory verbal hallucination (AVH), a Memory Based AVH or a Hypervigilance AVH

4. Inner Speech. This module provides psycho-education about the evidence that voice-hearing involves people not recognising their own inner speech. An individual understanding or formulation of voice-hearing experiences is co-produced and then targeted coping strategies and behavioural experiments are employed, such as means of interrupting and manipulating inner speech via singing or humming.

5. Memory-Based. This module provides psycho-education about how memories from trauma are more likely to be experienced as intrusive memories without contextual cues, and can therefore be experienced as belonging to the here and now. An individual formulation of how the memory may be experienced as a voice is followed by coping strategies and behavioural experiments that help people manage and reframe difficult memories.

6. Hypervigilance. This module provides psycho-education about how our brain uses prediction to interpret the world and manage the amount of sensory data received. If people are expecting threatening stimuli they may struggle to scrutinise poor quality sensory data and rely more heavily on predictions, whilst adopting a 'better safe than sorry' decision bias. These factors all make an individual more likely to hearing expected speech when it is absent. An individual formulation of how the hypervigilance hallucination occurred is developed and then targeted coping strategies and behavioural experiments are employed (such as reducing arousal and stress when under threat).

7. Seeing Visions. This module draws on these other modules, and explains how our visual perceptual system can lead to mistaken perceptions, for example how easily we see faces in clouds. An individual formulation and treatment plan is then developed that normalises the experience and addresses the key cause of distress and then targeted coping strategies and behavioural experiments are employed (such as training oneself to switch attention to and from visions).

8. Sleep. This module provides psycho-education and treatment strategies about sleep, which is often a key factor in all types of unusual sensory experiences.

6.2 Schedule of intervention

CPNs will use the manual in 4-6 therapy sessions. The number of sessions is based on previous work with other groups, but the clinician can choose to use the manual for more sessions, if they deem

necessary. This design will ensure that participants receive adequate exposure to the manual in therapy sessions for us to determine its acceptability.

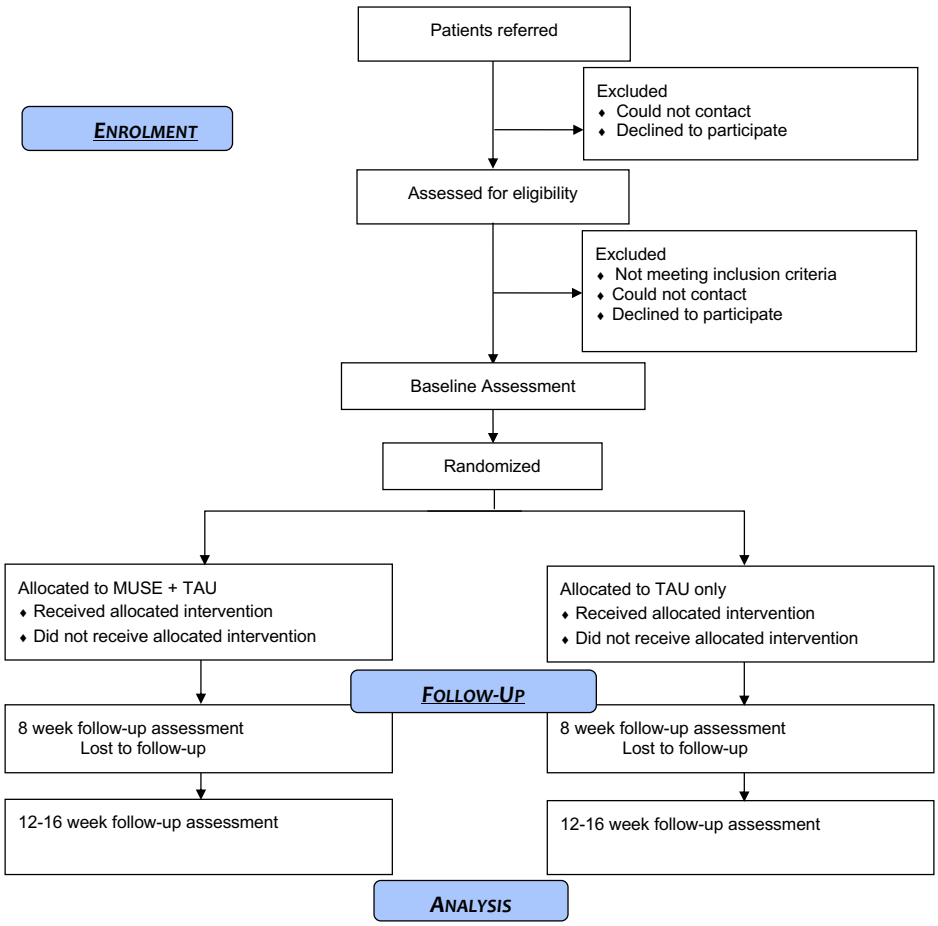
Criteria for discontinuing or modifying allocated interventions can be found in the trial protocol.

6.3 Participant timeline

Independent assessors (RWs) blind to treatment group will conduct all eligibility and research assessments. Following providing written informed consent to participate and completing the baseline assessment, eligible participants will be randomised within two working days. They will be contacted with the outcome of the randomisation within a week by the trial coordinator. Therapy will last up to 2 months in the intervention group. Participants will remain enrolled in the study for three to four months in total.

6.4 Study Flow Charts

Figure 1. Trial flow diagram



6.5 Inclusion Criteria

- Be in contact with Early Intervention in Psychosis (EIP) services
- Have an identified CPN acting as their care coordinator
- Meet ICD-11 criteria for schizophrenia, schizoaffective disorder or entry criteria for an EIP service
- Have a history of auditory hallucinations for at least four weeks
- Be aged 16 and above
- Consider their hallucinations as a main difficulty, and would like to receive an intervention specifically for hallucinations.
- Have the capacity to provide informed consent.
- Be judged by their clinician to be clinically stable for the preceding 4 weeks.
- Both individuals on antipsychotic treatment, and those who decline to take medication, will be included, as long as no medication changes have occurred in previous 1m (i.e., having started or stopped antipsychotic medication, or a switch to or from Clozapine);

6.6 Exclusion Criteria

- Hallucinations/psychosis with a known biological basis
- Insufficient command of English to complete the study procedures;
- Intellectual disability, or severe cognitive dysfunction affecting the ability to provide fully informed consent to participate;
- A primary diagnosis of substance misuse/dependency;
- Currently engaged in CBTp or received CBTp in the past 6 months

7 Randomisation and Blinding

Randomisation to the two groups will be undertaken using the web-based sealed envelope randomisation service. Randomisation will be in the ratio 1:1 to the two groups and will be stratified by site. Randomisation (at the individual level) will be independent and concealed, using permuted blocks of random size. The independent sealed envelope randomisation system will ensure blinding of the relevant members of the team. Outcome assessors will be blind. Clinicians, therapists and participants will be unblind. Trial statistician will be blind.

8 Study Variables

8.1 Demographic variables

- Age at randomisation
- Gender at birth (Male/Female)
- Ethnicity (White – Caucasian, Asian, Black, Middle-Eastern, Mixed-race, Other: Any)

8.2 Baseline variables and follow ups measures at 2 and 3 month post randomisation

In this section we describe outcome measures. Each outcome measure was scored according to standard methods. Please see appendix 1 for details. For scoring, except for SF-36v2, if there are more than 10% of items missing for a measure for a participant, the associated score will be considered as missing. When missing rate is less than 10%, missing items will be replaced by the median of non-missing item values across all other participants.

Psychotic Symptom Rating Scales (PSYRATS, (Haddock et al., 1999)) is a clinician administered semi-structured interview of hallucinations (such as amount/intensity of distress). It includes subscales measuring voice-related distress, frequency, attribution and loudness. It separately assesses delusions. We will report total score, total scores from each subscale, and total scores from assessment of delusions (See (Woodward et al., 2014)). Higher scores indicate worse symptoms.

Hamilton Program for Schizophrenia Voices Questionnaire (HPSVQ; (Van Lieshout & Goldberg, n.d.)) is a patient-reported questionnaire on auditory hallucinations. It has a subscale on voice impact. We will report both total score and score on the voice impact subscale. Higher scores indicate worse symptoms.

The Multimodal hallucination interview will be used to determine the number of unusual sensory modalities reported and will record presence of Auditory, Visual, Somatic, Olfactory and sense of presence. We will also calculate this with sensed presence removed, meaning score will be between 1-4. Higher scores indicate more modalities of hallucinations.

The short Depression, Anxiety and Stress Scales (DASS; (Lovibond & Lovibond, 1995)) is a 21 item self -report questionnaire designed to assess symptoms of anxiety, depression and stress. It has subscales for stress, anxiety and depression. We will report total score and scores for subscales. Higher scores indicate more negative affect and distress.

Questionnaire for the Process of Recovery QPR (Neil et al., 2009) is a user-defined measure, assessing subjective recovery in intrapersonal and interpersonal functioning. Our analysis will report total score for the QPR. Higher scores are indicative of recovery.

The CHoice of Outcome In Cbt for psychosEs (CHOICE; (Greenwood et al., 2010)), is a 21 item service-user developed questionnaire to evaluate outcomes for people with psychosis and assess therapy-related goals. We will use the short version in which scores from 12 questions are summed, and convert all questions so that higher scores indicate greater wellbeing.

SF-36 (Ware Jr & Sherbourne, 1992). The SF36 measures general activity and wellbeing in the past week. Subscales measure physical functioning, physical role, bodily pain, general health,

vitality, social functioning, role-emotional, mental health and health transition. We will measure each subscale, and summaries of mental and physical components. Higher scores represent a more favourable health state.

Investigating Choice Experiments Capability Measure for Adults (Flynn et al., 2015) consists of five domains (Secure, Support, Independence, Achievement, Enjoyment). We will report total score. Higher scores correspond to better health.

EQ-5d (Herdman et al., 2011) consists of questions about five domains (mobility, self care, usual activities, pain/discomfort, anxiety and depression) and a rating of current perception of health. We will consider each domain separately rather than a total score. Higher scores indicate worse perceived health.

Satisfaction with Therapy and Therapist Scale (Oei & Green, 2008) is a short scale assessing overall acceptability, with subscales assessing satisfaction with therapy and satisfaction with therapist. Higher scores represent greater satisfaction with therapy and the therapist.

Working Alliance inventory (Horvath & Greenberg, 1989) short form (Hatcher & Gillaspy, 2006) is used to assess alliance. Subscales assess agreement on the therapeutic task, the bond and agreement on goals. Higher scores on total score and subscales indicate greater therapeutic relationship.

Service use measure uses a clinical record form completed after follow up determining utilisation of health service. Adverse events will also be recorded separately.

8.3 Pseudo-primary outcome (PPO) and potential primary outcomes

Our pseudo-primary outcome is PSYRATS total score at end of treatment.

Our potential primary outcomes for an ongoing study are

- PSYRATS total score at follow up
- Hamilton total score at end of treatment
- PSYRATS voice impact subscale score at follow up

8.4 Overall screening rates

- Referral rate will not be used in the SAP as we relied on care coordinators to refer participants and there was not a general process to inform all service users in a team

about the study, therefore, we do not know if the referrals are representative of the broader pool of potential participants. Therefore, this is not an appropriate analysis.

- Recruitment rate will be calculated from the number referred and the number who agreed to be screened for assessment.
- Allocation compliance rate will be those who attended one or more session of MUSE if allocated to the condition.

For details of the overall screening rates see section

8.5 Derived variables

8.6 Schedule of Events:

Measure	Baseline	End of treatment	Follow up
PSYRATS • Total* • Voice-related distress* • Frequency • Attribution • Loudness • Delusions	x	x	x
Hamilton • Total* • Voice impact*	x	x	x
Multimodal hallucination interview • Total • Total excluding sensed presence	x	x	x
DASS • Total* • Stress • Anxiety • Depression	x	x	x
QPR • Total*	x	x	x
CHOICE • Total, short version*	x	x	x
SF-36 • Mental component summary* • Physical component summary* • Physical functioning	x	x	x

• Physical role • Bodily pain • General health • Vitality • Social functioning • Role-emotional • Mental health • Health transition			
ICECAP • Total*	x	x	x
EQ-5D • Mobility • Self care • Usual activities • Pain/discomfort • Anxiety • Depression	x	x	x
Service use measure			x
Working alliance			MUSE only
Satisfaction with Therapy and Therapist Scale			MUSE only

*Table 1: Timing of study measurements. Items marked * are included in calculations of effect size estimates.*

9 Sample Size

A formal sample size calculation has not been performed for this feasibility study. Our goal is not to assess treatment effectiveness but to establish whether we can undertake future pilot and definitive studies to address effectiveness. We have a target of recruiting 80 individuals over the recruitment period. Based on past research of psychological therapy conducted in the North East (Morrison et al., 2018; Thomson et al., 2017) and similar brief interventions (Foster et al., 2010) we have estimated attrition of 12.5% meaning approximately 70 people will complete the study. Guidance on external pilot studies indicates that samples of 35 per arm or more give a reliable estimate of the standard deviation of the outcome measure (Moore et al., 2011); however, the aim of the present study is not primarily to generate parameter estimates for a full trial, but to establish MUSE's acceptability and feasibility.

10 Statistical plans

10.1 Covariates and outcomes

We consider the outcomes marked * in **Error! Reference source not found.** (details in section 8.2) at both 8 weeks (post-treatment) and 12-16 weeks (follow-up).

In analyses including covariate adjustment, the following covariates will be included: age, sex, multimodal hallucination interview score (excluding sensed presence), baseline duration of auditory hallucinations, length of time engaged in service, site, and PSYRATS delusions score.

10.2 Analyses

10.2.1 Assessment of effect of treatment on PPO, no covariates other than site

We will assess impact of treatment against the potential primary outcome (PSYRATS total at 8 weeks) with site as a covariate and with no additional covariates. Site will be modelled as a fixed effect, as it has only two levels in this study. This will be on the basis of intention-to-treat and analysers will be blinded to treatment group identity.

We will assume a rate of unit dropouts independent of treatment status, and hence remove dropouts from the analysis. We will use a linear mixed model with baseline and follow-up PSYRATs score included as detailed in appendix 2.

We will estimate the coefficient of the treatment indicator and the standard error/confidence interval of this estimate. We will estimate the minimum sample size for a definitive trial to detect an effect of this estimated size at a type-1 error rate of 5% with 80% power.

Under the assumption that the PPO and all potential primary outcomes have equal true effect sizes, this will constitute an asymptotically unbiased estimate of the minimum sample size necessary to detect this common effect (see appendix 2).

10.2.2 Assessment of effect of treatment on PPO, covariates included

We will assess impact of treatment against the potential primary outcome (PSYRATS total) with covariates included as indicated above. All covariates will be modelled as fixed effects. We will use a linear mixed model as specified in appendix 2.

We will estimate coefficients of treatment and covariate indicators and report these with confidence intervals.

10.2.3 Within-treatment analysis of number of treatments

We will assess the effect of the number of MUSE treatments on the PPO with covariates other than site not included, using a linear mixed effects model as detailed in appendix 2. This analysis will use only individuals in the treatment group and dropped-out individuals will be excluded. The number of completed MUSE treatments will be modelled as a fixed effect.

10.2.4 Assessment of missingness, adherence, and other variables

We will summarise measures of missingness and adherence throughout the study, using confidence intervals where appropriate. In particular we will assess whether there is evidence of non-random dropout by comparing proportions of dropouts and missing outcomes in each treatment group.

We will repeat analysis 1 on the per-protocol population and potentially the as-treated and safety populations, and assess whether results change substantially.

We will assess differences between study sites for covariates and outcomes above.

We will assess the impact of COVID on the study, as indicated below.

10.2.5 Assessment of effect of treatment against other outcomes, no covariates other than site

We will repeat the procedures of the first analysis above substituting the other indicated outcomes for the PPO. Again, coefficients of the treatment indicator will be reported to summarise effect sizes, and confidence intervals will be computed.

10.2.6 Assessment of effect of treatment against other outcomes, covariates included

We will repeat the procedures of the second analysis above substituting the other indicated outcomes for the PPO and including covariates as fixed effects. Effect sizes will be summarised as above.

10.2.7 Assessment of potential primary outcomes

We will assess the suitability of the potential primary outcomes. This will include consideration of adherence, acceptability, effect size, and similarity to the PPO. Similarity to the PPO will be assessed on the basis of Spearman (rank) correlation between scores.

10.2.8 Planning of definitive trial

We will plan a potential definitive trial, after assessing whether the criteria for progression to such a trial are met (see later section). The effect size for the PPO and corresponding sample size will be considered in light of potential biases (see appendix 2).

A primary outcome will be selected. Adherence, missingness, and site differences will be considered. We will establish whether covariates should be included in the analysis of the primary outcome.

10.3 Reporting and blinding

All outcomes will be reported in full in any publication. We will conduct these analyses in order. Analysis staff will be unblinded to treatment group after analysis 1 has been completed.

11 General Considerations

11.1 Timing of Analyses

The total duration of this trial was anticipated to be 19 months. Participants will be assessed on a range of measures before and after treatment (at 2 months) and about a month later (at 3 months from the start of the treatment). The final 4-5 months will be comprised of final data analysis, report writing and dissemination of research findings. The final analysis will be performed at the end of the trial after the data cleaning, data lock and after this Statistical

Analysis Plan is approved by the Chief Investigator, trial statisticians and Chair of Data Monitoring Committee.

11.2 Analysis Populations

11.2.1 Full Analysis Population

The full analysis population includes all participants who were consented and randomised under the intention-to-treat principle and who agreed that their data can be analysed following completion or withdrawal from the trial. All analyses will be based on the full analysis population following the intention-to-treat principle.

11.2.2 Per Protocol Population

The per-protocol population will be used mainly for sensitivity and exploratory analyses and it will be considered for all outcomes. This will be descriptive. Potential future primary outcome measures include PSYRATS total, PSYRATS H-Distress, Hamilton total and Hamilton distress, and QPR and these will be considered specifically in a per protocol report. It includes all participants who received the treatment group that attended 4 or more sessions, completed end of treatment (2 month) follow-up and did not deviate from the protocol as agreed with the TSC. The per-protocol population will usually be smaller than the full analysis population due to the deletion of participants who violated the trial protocol.

11.2.3 As treated population

It is not anticipated that there will be a large number of people who receive MUSE when allocated to TAU. We are investigating contamination of the TAU group specifically using the Tidier checklist (Hoffman et al., 2014). If found that there is a considerable amount of MUSE offered in TAU (>50% of the TAU group) then the as treated analysis will be undertaken. The as-treated population will also be used mainly for sensitivity and exploratory analyses and it will be considered for all outcomes. All participants will be analysed according to the actual intervention they received instead of the intervention to which they were randomised. This will be most relevant to participants who were randomised to an intervention but received the other one post-randomisation. The as-treated population will usually be the same as the full analysis population, except for missing data.

11.2.4 Safety Population

All subjects who received any study treatment (including control) and are not dropped out due to serious adverse events post randomisation. Safety population will be considered for number of Serious adverse events (SAEs) and adverse events (AEs) in the outcome analysis.

11.3 Covariates and Subgroups

Sub-group analysis will be limited as this is a feasibility study. However, we will consider if there are differences in baseline characteristics between sites (age, gender, PSYRATS total score), and duration of voices (derived from the demographics information) and duration of time in the EIP team. For consideration of treatment effects, we will consider if duration of

voices, time in EIP and PSYRATs delusions, and number of hallucination (derived from the Multimodal Hallucination Interview) scores moderates treatment outcome.

11.4 Interim Analyses

There will be no interim analysis planned for this trial.

11.5 Protocol Deviations

Protocol deviations will be recorded and include issues like delayed randomisations, provision of therapy by a psychological therapist rather than a care coordinator, delayed follow up assessments etc. These will be reported descriptively and not subject to formal analysis.

11.6 Treatment Compliance

Treatment compliance has been agreed with the TMM and the PPI group and stated in the trial protocol as 4 or more sessions of MUSE. This information will be presented descriptively. We will also present information on the modules that have been used by the therapists.

11.7 Criteria for proceeding to a full trial

We will use criteria for assessing study success and identifying feasibility factors required for delivering the definitive study (Malterud et al., 2016) and follow a systematic process for decision making after pilot and feasibility trials (ADePT, (Bugge et al., 2013)) which helps identify criteria used to go to a full trial. These criteria have been developed with the LEAP to help determine if a full trial is warranted. We will likely use criteria on participant recruitment, adherence with the intervention, and retention at follow-up to assess the trial (as set out in ADePT), plus data on uptake, retention of participants, intervention fidelity and acceptability. This will use both quantitative and qualitative data derived from the study. The progression criteria will be divided into three categories (green, red and amber; (Malterud et al., 2016)). Areas that are amenable to change before a pilot trial will be investigated and solutions discussed with the patient LEAP for acceptability. This will help consider if a full trial is timely, necessary, and deliverable.

Following discussion with the LEAP and reviewing other research conducted with similar populations the following criteria have been set and agreed. The PPI group identified three criteria, indicated below in the table, that they felt were particularly important to consider in deciding on the feasibility of a future trial. These focussed on whether the trial could recruit to target in a timely manner, whether people remain in the study, and whether treatment was delivered in an adequate amount in a timely manner.

	Red Do not proceed	Amber Consider revisions	Green Proceed to trial
*Rate of referral	N/A	N/A	N/A
Recruitment rate (of the number referred, what % agreed to be screened)	Under 60%	60-79%	80% or more
Allocation compliance (number of individuals attending one session of MUSE or TAU)	Under 60% of target met	60-79% of target met	80% or more
Recruitment target n=80**	Under 60% of target met	60-79% of target met	80% of target met
Retention at 8-10 weeks**	Under 60% retained	60-87.4% retained	87.5% retained
Retention at 12-16 weeks	Under 60% retained	60-87.4% retained	87.5% retained
Treatment compliance defined as receiving 4 or more sessions **	Under 60%	60-80%	80% received 4 or more sessions

*following consultation with the TMM and PPI group this has been excluded as a success criteria. Given we did not proactively recruit or attempt to contact all the service users in EIP we were reliant on contact via care coordinators, therefore, we are unable to estimate the value.

** identified as a key criteria by the PPI group

11.8 Impact of Covid

The trial was run during the pandemic. It began recruitment in July 2021 until May 2022. In January 2022 there was a substantial increase in Covid, attributed to the Omicron variant. There were significant impacts in the NHS and the trial lost 7 of the 15 recruiting sites owing to service pressures. The impact was felt in terms of reduced referrals, and reduced therapist capacity. Moreover, Covid affected service users, their families as it did the therapists in the trial and isolation rules meant there could be delays in starting/delivering treatment. This impacted on the time that it took to deliver the therapy, and led to delays in provision of the MUSE treatment. This effected when the follow up appointments could be provided as in many cases the person had not had time to receive treatment sessions. So, from January onwards typically follow ups were delayed to 10 or more weeks rather than 8 and so the subsequent follow up was at 14-16 weeks rather than at 12. We had randomised around 40 of the participants by December 2021, so we will examine the impact of COVID on the timing of the assessments, and the number of sessions provided, and explore if it has any impact on outcome (PSYRATs total, AH Distress, and Hamilton and QPR) from a descriptive perspective.

12 Efficacy Analyses

Descriptive statistics within each randomised group will be presented for demographic data, for variables in section 8.2 at each follow up time point and for overall screening rates in section 8.4. These will include frequency and percentages for binary and categorical variables, and means and standard deviations, or medians with interquartile range (QR), for continuous variables, along with minimum and maximum values, and frequency and percentages of missing values. There will be no tests of statistical significance or confidence intervals for differences between randomised groups on any baseline variable, nor will attempts at inference be made on the basis of confidence intervals (see appendix 2).

12.1 Missing Data

We will provide rate of missing data, which will be calculated as the proportion of participant with available data in PSYRATS total at baseline, 2 and 3 months. Using cross-tabulation, we would assess whether participants per site and in the treatment groups are equally likely to report missing data.

13 Safety Analyses

All analyses will be in line with the safety reporting and safety criteria in the protocol. Adverse (AE) Events and Serious Adverse Events (Adverse Events which meet the criteria for seriousness) will be captured for the participants. AEs and SAEs will be tabulated per trial arms and the action taken, outcome and causality in the opinion of the investigator will be reported using frequency tables.

14 Reporting Conventions

Means, standard deviation, and any other statistics other than quantiles, will be reported to two decimal places greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters not on the same scale as raw observations (e.g. regression coefficients and confidence intervals) will be reported to 2 decimal places.

15 Technical Details

All analysis will be done on a using the statistical programming language R. The version will be recorded prior to the analysis. We will use the LMER package to fit linear mixed models.

The population to be used in a table or figure will be explicitly set at the start of a block of code that computes the output, ideally by looking up the population from the table of tables. Any code will have

- The date and time included
- The name of the code file that produced the analysis
- The author
- The date and time of writing
- References to inputs and outputs
- Reference to any parent code file that runs the child code file

16 References

Andrew, A., Knapp, M., McCrone, P., Parsonage, M., & Trachtenberg, M. (2012). Effective interventions in schizophrenia: The economic case. *A Report Prepared for the Schizophrenia Commission. London: Rethink Mental Illness*.

Bugge, C., Williams, B., Hagen, S., Logan, J., Glazener, C., Pringle, S., & Sinclair, L. (2013). A process for Decision-making after Pilot and feasibility Trials (ADePT): Development following a feasibility study of a complex intervention for pelvic organ prolapse. *Trials*, *14*(1), 1–13.

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., & Lancaster, G. A. (2016). CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. *Bmj*, *355*.

Flynn, T. N., Huynh, E., Peters, T. J., Al-Janabi, H., Clemens, S., Moody, A., & Coast, J. (2015). Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. *Health Economics*, *24*(3), 258–269.

Foster, C., Startup, H., Potts, L., & Freeman, D. (2010). A randomised controlled trial of a worry intervention for individuals with persistent persecutory delusions. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*(1), 45–51.

Greenwood, K. E., Sweeney, A., Williams, S., Garety, P., Kuipers, E., Scott, J., & Peters, E. (2010). CHoice of Outcome In Cbt for psychosEs (CHOICE): The development of a new service user–led outcome measure of CBT for psychosis. *Schizophrenia Bulletin*, *36*(1), 126–135.

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: The psychotic symptom rating scales (PSYRATS). *Psychological Medicine*, *29*(4), 879–889.

Hatcher, R. L., & Gillaspy, J. A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. *Psychotherapy Research*, *16*(1), 12–25.

Herdman, M., Gudex, C., Lloyd, A., Janssen, M. F., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, *20*(10), 1727–1736.

Hjorthøj, C., Stürup, A. E., McGrath, J. J., & Nordentoft, M. (2017). Years of potential life lost and life expectancy in schizophrenia: A systematic review and meta-analysis. *The Lancet Psychiatry*, *4*(4), 295–301.

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, *36*(2), 223.

Kennedy, J. L., Altar, C. A., Taylor, D. L., Degtiar, I., & Hornberger, J. C. (2014). The social and economic burden of treatment-resistant schizophrenia: A systematic literature review. *International Clinical Psychopharmacology*, *29*(2), 63–76.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343.

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, *26*(13), 1753–1760.

Moore, C. G., Carter, R. E., Nietert, P. J., & Stewart, P. W. (2011). Recommendations for planning pilot studies in clinical and translational research. *Clinical and Translational Science*, *4*(5), 332–337.

Morrison, A. P., Pyle, M., Gumley, A., Schwannauer, M., Turkington, D., MacLennan, G., Norrie, J., Hudson, J., Bowe, S. E., & French, P. (2018). Cognitive behavioural therapy in clozapine-resistant schizophrenia (FOCUS): An assessor-blinded, randomised controlled trial. *The Lancet Psychiatry*, *5*(8), 633–643.

Neil, S. T., Kilbride, M., Pitt, L., Nothard, S., Welford, M., Sellwood, W., & Morrison, A. P. (2009). The questionnaire about the process of recovery (QPR): A measurement tool developed in collaboration with service users. *Psychosis*, *1*(2), 145–155.

NICE. (2014). *Psychosis and schizophrenia in adults: Prevention and management*. NICE.

Oei, T. P., & Green, A. L. (2008). The Satisfaction With Therapy and Therapist Scale–Revised (STTS-R) for group psychotherapy: Psychometric properties and confirmatory factor analysis. *Professional Psychology: Research and Practice*, *39*(4), 435.

Thomson, C., Wilson, R., Collerton, D., Freeston, M., & Dudley, R. (2017). Cognitive behavioural therapy for visual hallucinations: An investigation using a single-case experimental design. *The Cognitive Behaviour Therapist*, *10*.

Van Lieshout, R. J., & Goldberg, J. O. (n.d.). Hamilton Program for Schizophrenia Voices Questionnaire. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*.

Ware Jr, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, 473–483.

Woodward, T. S., Jung, K., Hwang, H., Yin, J., Taylor, L., Menon, M., Peters, E., Kuipers, E., Waters, F., & Lecomte, T. (2014). Symptom dimensions of the psychotic symptom rating scales in psychosis: A multisite study. *Schizophrenia Bulletin*, *40*(Suppl_4), S265–S274.