# Assessing the efficacy and safety of STOP (successful treatment for paranoia)—an app-based cognitive bias modification therapy for paranoia: a randomised clinical trial protocol

Jenny Yiend[1*†] , Rayan Taher[1†], Carolina Fialho[1], Chloe Hampshire[2], Che-Wei Hsu[1,7,8], Thomas Kabir[3], Jeroen Keppens[4], Philip McGuire[3], Elias Mouchlianitis[1], Emmanuelle Peters[1,5], Tanya Ricci[1], Sukhwinder Shergill[6], Daniel Stahl[1], George Vamvakas[1], Pamela Jacobsen[2], the MPIT and Avegen

## Abstract

**Background** Paranoia, the belief that you are at risk of significant physical or emotional harm from others, is a common difficulty, which causes significant distress and impairment to daily functioning, including in psychosis-spectrum disorders. According to cognitive models of psychosis, paranoia may be partly maintained by cognitive processes, including interpretation biases. Cognitive bias modification for paranoia (CBM-pa) is an intervention targeting the bias towards interpreting ambiguous social scenarios in a way that is personally threatening. This study aims to test the efficacy and safety of a mobile app version of CBM-pa, called STOP (successful treatment of paranoia).

**Methods** The STOP study is a double-blind, superiority, three-arm randomised controlled trial (RCT). People are eligible for the trial if they experience persistent, distressing paranoia, as assessed by the Positive and Negative Syndrome Scales, and show evidence of an interpretation bias towards threat on standardised assessments. Participants are randomised to either STOP (two groups: 6- or 12-session dose) or text-reading control (12 sessions). Treatment as usual will continue for all participants. Sessions are completed weekly and last around 40 min. STOP is completely self-administered with no therapist assistance. STOP involves reading ambiguous social scenarios, all of which could be interpreted in a paranoid way. In each scenario, participants are prompted to consider more helpful alternatives by completing a word and answering a question. Participants are assessed at baseline, after each session, and at 6, 12, 18 and 24 weeks post-randomisation. The primary outcome is the self-reported severity of paranoid symptoms at 24 weeks, measured using the Paranoia Scale. The target sample size is 273 which is powered to detect a 15% symptom reduction on the primary outcome. The secondary outcomes are standardized measures of depression, anxiety and recovery and measures of interpretation bias. Safety is a primary outcome and measured by the Negative Effects Questionnaire and a checklist of adverse events completed fortnightly with researchers. The trial is conducted with the help of a Lived Experience Advisory Panel.

**Discussion** This study will assess STOP's efficacy and safety. STOP has the potential to be an accessible intervention to complement other treatments for any conditions that involve significant paranoia.

---

†Jenny Yiend and Rayan Taher are joint first authors.

*Correspondence:
Jenny Yiend
jenny.yiend@kcl.ac.uk
Full list of author information is available at the end of the article

## Administrative information

| | |
|---|---|
| Title {1} | STOP—Successful Treatment of Paranoia: Replacing harmful paranoid thoughts with better alternatives |
| Trial registration {2a and 2b} | ISRCTN17754650 |
| Protocol version {3} | 3.2 |
| Funding {4} | This research was funded by the Medical Research Council (MRC) Biomedical Catalyst: Developmental Pathway Funding Scheme (DPFS), MRC Reference: MR/V027484/1 |
| Author details {5a} | Jenny Yiend[1], Rayan Taher[1], Carolina Fialho[1], Chloe Hampshire[2], Che-Wei Hsu[1,12,13], Thomas Kabir[3], Jeroen Keppens[4], Philip McGuire[1], Elias Mouchlianitis[1], Emmanuelle Peters[5], Tanya Ricci[1], Sukhwinder Shergill [1,7], Daniel Stahl[8], George Vamvakas[9], Pamela Jacobsen[2], the MPIT[10] and Avegen[11] *[1]Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, [2]Department of Psychology, University of Bath, Bath, UK [3]The McPin Foundation, London, UK [4]Department of Informatics, King's College London, London, UK [5]Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK [6]South London and Maudsley NHS Foundation Trust, London UK [7]Kent and Medway Medical School, Canterbury, Kent, United Kingdom* [8]Department of Biostatistics and Health Informatics, *Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK [9]Clinical Trials Unit, King's College London, London, UK [10]McPin Public Involvement Team includes, in alphabetical order: Al Richards, Alex Kenny, Edmund Brooks, Emily Curtis, and Vanessa Pinfold [11]The Avegen group includes, in alphabetical order: Nandita Kurup, Neeraj Apte and Neha Gupta [12]Department of Psychological Medicine, Dunedin School of Medicine, University of Otago, NZ [13]College of Community Development and Personal Wellbeing, Otago Polytechnic, NZ* |

| | |
|---|---|
| Name and contact information for the trial sponsor {5b} | **Lead Sponsor** Professor Reza Razavi Vice President & Vice Principal (Research) King's College London Room 5.31, James Clerk Maxwell Building 57 Waterloo Road London SE1 8WA Tel:+44 (0)207 8,483,224 Email: reza.razavi@kcl.ac.uk **Co-Sponsor** Mr Dunstan Nicol-Wilson South London and Maudsley NHS Foundation Trust R&D Department Room W1.08 Institute of Psychiatry, Psychology & Neuroscience (IoPPN) De Crespigny Park London SE5 8AF Tel:+44 (0)207 8,480,339 Email: slam-ioppn.research@kcl.ac.uk |
| Role of sponsor {5c} | Neither the Funder nor the Sponsor will have a role in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication |

## Introduction

### Background and rationale {6a}

Paranoia refers to the belief that other people are deliberately intending to cause harm to you in some way, including emotional (e.g. humiliation) and physical harm (e.g. your life being in danger). It also includes the fear of coming to harm from physical objects and situations (e.g. security cameras). Psychological models of paranoia suggest that it occurs on a continuum. Paranoid thinking in the general population is thought to have a hierarchical structure, ranging from mild social evaluative concerns (e.g. fears of rejection) to severe threats (e.g. significant personal harm) [1–5]. Persecutory delusions fall at the extreme point on the continuum of paranoid belief. As such, they are associated with more distress than other types of delusion [6], are most likely to elicit a behavioural response to cope with the experience [7] and are a strong predictor of hospitalization [8]. Over

one-third of UK psychiatric patients suffer from persecutory delusions, often appearing in a range of psychopathologies, including depression [9], bipolar disorder [10], posttraumatic stress disorder [11], anxiety [12], and with the highest prevalence and greatest intensity in psychosis [13].

National Institute for Health and Care Excellence (NICE) clinical guidelines for psychosis and schizophrenia recommend that service users are offered cognitive behavioural therapy (CBT). However, some estimates suggest that CBT is received by as little as 24% of those who could benefit and it has shown only moderate effect sizes for delusions [14–16]. Furthermore, a significant proportion of patients suffering from persecutory delusions continue to experience distressing symptoms following psychological treatment [17, 18]. Thus, there is a need for novel, accessible, scalable interventions for paranoia, either as standalone treatments or as adjuncts to boost existing therapies. This is also important for widening access to help and support for people who may experience paranoia which causes distress and impairment, but who do not present to mental health services, or who may not want to take up the offer of CBT delivered by a therapist.

Recent thinking in the treatment for delusions emphasises briefer, targeted interventions, with a focus on putative causal factors such as cognitive biases [19, 20]. Here we use the term 'cognitive bias' in its narrower sense to refer to the selective processing of information that matches the core content of the pathology of a disorder [5]. For example, individuals with persecutory delusions may consistently interpret emotionally ambiguous information in one unhelpful direction (i.e. interpretation bias). For example, two people laughing together at a bus stop may be taken to mean 'they are laughing at me/trying to humiliate me'. Repeated paranoid interpretations can produce an increased perception of danger of harm to the self, which can reinforce paranoid beliefs, maintain symptoms and increase distress. This, in turn, triggers even more biased processing in a self-perpetuating cycle. In theoretical models of psychosis, this cycle is proposed as contributing to the formation and maintenance of positive symptoms such as delusions [6].

For these reasons, brief psychological treatments that target cognitive biases directly, such as cognitive bias modification (CBM), have been developed [21]. CBM is a class of theory-driven treatments that use laboratory-derived, repetitive word tasks to manipulate biases and promote more adaptive processing and have been used to ameliorate anxiety and depression [21]. CBM-pa (cognitive bias modification for paranoia) is a six-session computer desktop version of this class

of intervention specifically designed to target paranoid interpretations, i.e. it 'trains' the brain to process and interpret emotionally ambiguous situations in a more benign, non-paranoid way [22]. Feasibility testing of a six-session CBM-pa has shown promising results [23]. STOP (successful treatment of paranoia) is the successor to CBM-pa. STOP takes the form of a 12-session, self-administered mobile app with audio and graphic enhancements of its therapeutic content. The development of STOP was an iterative and thorough process that involved collaboration between individuals with lived experience of paranoia, clinical psychologists and researchers with full details having been published elsewhere [24].

The present paper presents the full protocol of the STOP randomised controlled trial (RCT) to assess its efficacy and safety for transdiagnostic use in people with distressing clinical levels of paranoia. The use of an active control (an identical app, but replacing the therapeutic text and tasks with neutral content) permits robust conclusions about the mechanism of action of any therapeutic effects.

### Objectives {7}
The primary objectives of the trial are as follows:

- To assess the efficacy of STOP.
  Efficacy is evaluated in two different doses (6 and 12 sessions) using a primary outcome (paranoia scale) measured at 24 weeks post-randomisation with a 12-session text reading control.
- To assess the safety of STOP.
  Safety is evaluated using the Negative Effects Questionnaire at 12 weeks and a fortnightly checklist of adverse events.

A secondary objective is to examine the dose–response relationship by performing a retrospective time-series dose–response modelling of change over time on key outcome measures.

### Trial design {8}
The trial is a double-blind, superiority, three-arm RCT recruiting individuals across England and Wales who experience clinical levels of paranoia as assessed by the Positive and Negative Syndrome Scales and show evidence of an interpretation bias towards threat on standardised assessments. Participants are randomly allocated to a 6-session intervention group, a 12-session intervention group or a 12-session text reading control. Treatment as usual (TAU) will continue for all

participants. Sessions are weekly and last around 40 min. Assessments are at baseline, after each session, and at 6, 12, 18 and 24 weeks post-randomisation. The trial is designed to run entirely remotely.

## Methods: participants, interventions and outcomes
### Study setting {9}
This study is conducted across two main recruitment hubs. The lead hub is the Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London. The second hub is the Department of Psychology, University of Bath. In addition, 20 NHS Trusts across England and Wales were recruited to the study through the Clinical Research Network (CRN) portfolio. The trial is conducted with the help of a lived experience advisory panel (LEAP).

### Eligibility criteria {10}
There are no eligibility criteria relating to current or past use of mental health services or diagnosis, as we aim to assess the use of STOP in a wide range of people experiencing paranoia, including those currently underserved by available sources of help and support.

*Inclusion criteria:*

1. Any clinically significant persecutory or paranoid symptoms, present for at least the preceding month. This is operationalised as a score of 3 ('mild') or more on item 6 of the Positive and Negative Syndrome Scale (PANSS) [25]. Item 6 measures 'suspiciousness/ persecution' defined as 'unrealistic or exaggerated ideas of persecution, as reflected in guardedness, a distrustful attitude, suspicious hyper-vigilance, or frank delusions that others mean one harm'.
2. Displaying an interpretation bias on the 8-item screening version of the similarity ratings task (SRT). The SRT screening task is a shortened (8-item) version of the SRT [26, 27]. Screening bias scores range from $+3$ to $-3$. Zero indicates no bias; plus scores indicate a bias favouring paranoid interpretations and minus scores indicate a bias favouring nonparanoid interpretations. A bias score of $-2$ or greater (i.e. $-2$ up to $+3$) is necessary to be selected for the study.
3. If on psychotropic medication, participants should be stable on that medication for at least the last 3 months and expected to be so for the study duration.
4. Age 18 years or older.
5. Capacity to consent: if in doubt, this is assessed using the Capacity Assessment Tool [28].

*Exclusion criteria:*

1. Severe cognitive impairment (e.g. being unable to validly complete one or more of the screening assessments)
2. Illiteracy or inability to understand written and spoken English.
3. Major current physical illness (e.g. cancer, heart disease, stroke, dementia).
4. Major substance or alcohol misuse, assessed by CAGE-AID Substance Abuse Screening Tool [29].
5. Currently receiving, or soon due to receive, a psychological intervention or having done so in the last 3 months. A psychological intervention, for the purpose of the study, is defined as five or more consecutive sessions with a trained therapist, either in an individual or group setting, for the purpose of alleviating symptoms arising from paranoia or related issues.
6. Currently taking part in any other interventional research study.
7. Extreme paranoia, i.e. scoring 7 (defined as 'Extreme— a network of systematised persecutory delusions denotes the patient's thinking, social relations and behaviour') on item 6 of the PANSS.
8. At high risk of suicide as indicated by the Columbia Suicide Severity Rating Scale (C-SSRS)—Screen Version [30, 31].

### Who will take informed consent? {26a}
Consent and screening are carried out by either.

i) Trained research assistants from the University of Bath and King's College London who are part of the main study team or
ii) Good Clinical Practice (GCP)-trained CRN or clinical staff at recruiting sites across England and Wales.

Capacity to consent is assumed in accordance with government legislation [32]. Where there is doubt, capacity is assessed during screening using a capacity assessment tool that is based on this legislation.

### Additional consent provisions for collection and use of participant data and biological specimens {26b}
We include an explicit item in our informed consent procedures explaining the process for fully anonymised data sharing via the national repository, UK Data Archive and data files deposited at the ISRCTN registry (https://www.biomedcentral.com/p/isrctn). Participants

who do not wish their anonymised data to be deposited can omit checking this box and/or strike out this item.

This trial does not involve collecting biological specimens for storage.

## Interventions

### Explanation for the choice of comparators {6b}

The study uses an active control condition to enable the evaluation of the effect of manipulating the hypothesised mechanism of action (bias reduction) directly, controlling for any non-specific effects (e.g. contact with researchers, interaction with digital equipment, trial participation, etc.) in a manner not possible with a waiting list control.

### Intervention description {11a}

This trial will aim to compare two interventions, the STOP mobile app and a text reading control. Both are described here.

### STOP intervention

Participants randomised to the intervention receive 40 training items per session on the STOP mobile app.

Participants read text inviting paranoid interpretations, then complete missing words and answer questions in a way that encourages alternative, adaptive beliefs about themselves and others (see Fig. 1).

Full details of the development of the STOP intervention have been published by Hsu and colleagues [24]. Here we provide an overview of key points. The item content, including scenarios, questions and responses, was derived from real-life examples generated by 18 people with lived experience of clinical paranoia. The academic team then revised the questions and answers into the standard format shown in Fig. 1 (three-line stories with a final word completion resolving ambiguity positively, followed by a one-sentence question to reinforce the positive resolution).

### Severity grading of sessions

The order of intervention sessions was informed by the paranoia hierarchy [1]. Clinician-delivered exposure therapy involves gradually exposing clients to increasingly triggering situations, meaning early sessions are less anxiety-provoking and more tolerable for patients. This
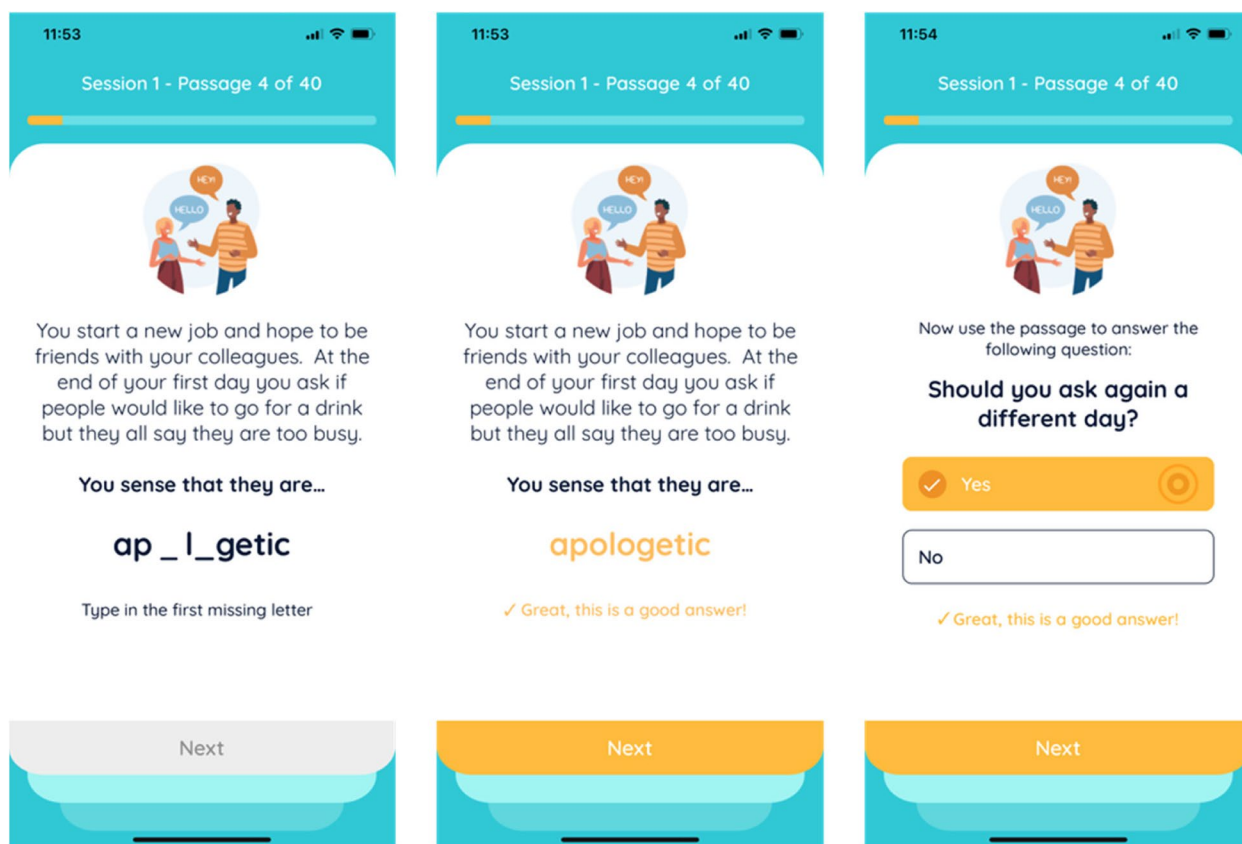
**Fig. 1** STOP training item example

allows habituation of the emotional response to occur gradually as the patient progresses through to more challenging therapeutic content and is usually more acceptable to patients, reducing the risk of dropout. Similarly, STOP's early sessions include less severe items, with a graded progression toward more challenging items in later sessions.

### Verb grading of sessions

Clinician-administered cognitive therapies traditionally employ an increasing 'drill-down' from more surface-level automatic thoughts, into more rigid rules and assumptions, and then finally towards deeper core beliefs. In order to reflect this in STOP, we selected specific verbs for the final sentence of each passage. In sessions 1–4, the following verbs are used to reflect automatic thoughts: 'think', 'imagine' or 'sense' (e.g., you *think* that this means…). In sessions 5–8, the following verbs are used to capture underlying assumptions: 'assume', 'presume' and 'suppose' (e.g. you *assume* that this means …). In sessions 9–12, core beliefs are reflected using the verbs 'believe', 'are sure' and 'know' (e.g. you *believe* that this means …). Appropriate verbs to capture each level were decided by consensus of three members of the research team with experience in paranoia, cognitive therapies and CBM methods.

In the six-session arm of STOP, users receive only the first 6 sessions of the full 12-session therapy. The content of the first six sessions is therefore less severe in terms of quantitative paranoia ratings and uses only the earlier verbs in the drill-down sequence.

### Text-reading control

Participants randomised to the active control condition receive 40 control items per session (approximately 40 min each) on the STOP mobile app. The experience is identical to the intervention condition except the item content that participants see omits the 'active ingredient' (resolution of a potentially triggering, but ambiguous, situation in a benign/non-paranoid manner. Instead, control participants read and respond to factual material or mundane everyday experiences (see Fig. 2).

### Interventions: modifications {11b}

The trial arm interventions cannot be modified for individual participants. Participants choosing to take
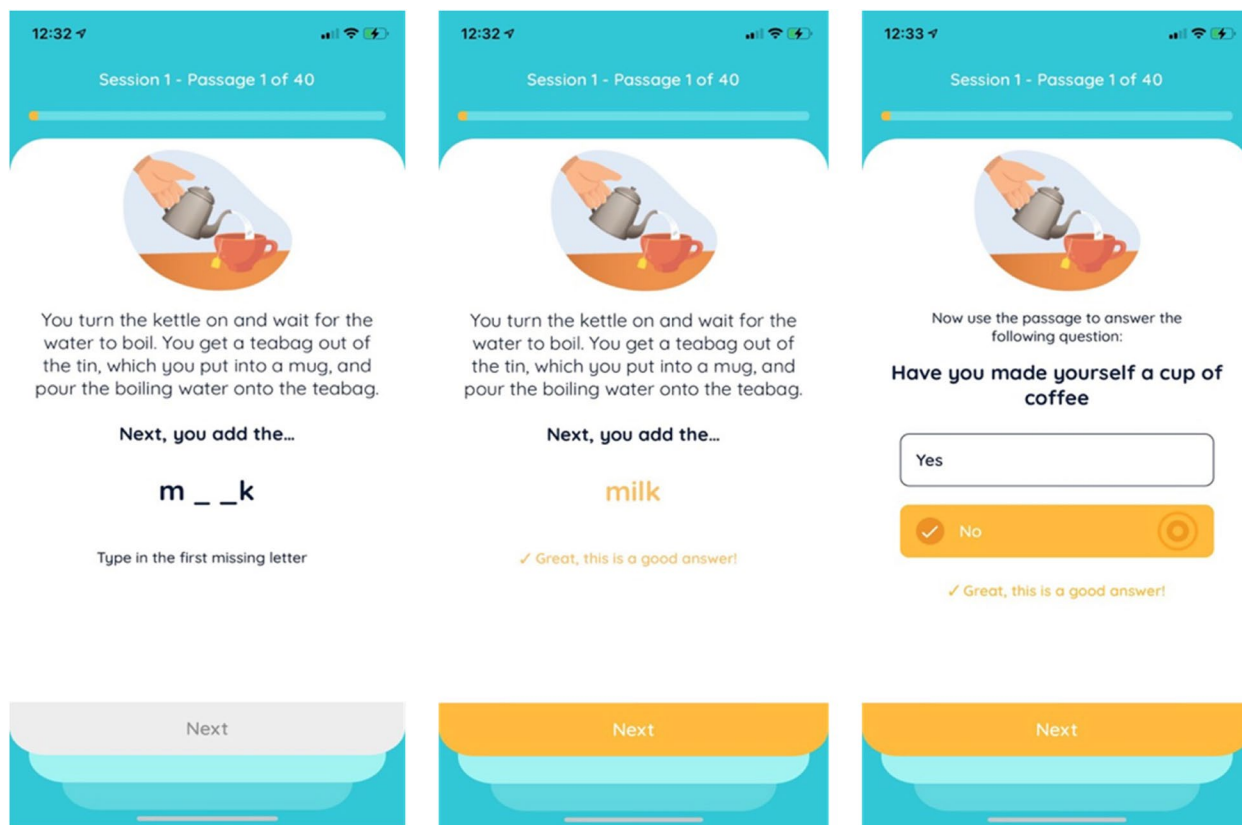
**Fig. 2** Control item example

part will follow the procedures as predefined for their allocated arm. However, participants have the right to withdraw from the study at any time for any reason. Participants may also be withdrawn from the trial if deemed appropriate by the trial steering committee (TSC) or chief investigator (CI), for example, due to adverse events (AEs), serious adverse events (SAEs), protocol violations, administrative or other reasons.

### Strategies to improve adherence to interventions {11c}

Adherence is a known challenge faced by online interventions. To mitigate this, STOP includes evidence-based gamification techniques including progress tracking and reward badges. During the trial, the following will be implemented to reduce drop-out:

1. When registering patients into the trial, researchers and participants schedule dates together for completing each session using the app's calendar. A (blinded) administrator portal tracks participants' progress. Missed or incomplete sessions or assessments trigger auto-alerts to the team for investigation and/or additional support.
2. Participants (intervention and control) receive a weekly or fortnightly phone call from researchers, following a pre-prepared script.
3. STOP and the text reading control condition include audio-visual enhancements (e.g. dynamic and static images and sounds) to increase participants' interest and engagement with the text they are reading.

Adherence to the intervention is defined in Table 1 and is extracted automatically by the app upon a participant's exit from the trial.

### Relevant concomitant care permitted or prohibited during the trial {11d}

Care permitted: all three conditions are conducted in addition to TAU, which consists of individualised combinations of medication and care coordination.

Care prohibited: none. Patients are not eligible for the trial if they are in receipt of a conflicting psychological therapy currently, within the previous 3 months, or are likely to start such a therapy for the duration of the trial. If a patient starts a therapy while in the trial this is recorded under TAU data collection; they would not be withdrawn from the trial.

### Provisions for post-trial care {30}

TAU will continue upon participants' exit from the trial.

### Outcomes {12}

The full assessment schedule is shown in Table 2.

### Primary outcome measure

The primary outcome is mean paranoid symptoms measured using the paranoia scale (PS) [33] at 24 weeks post-randomisation. The PS is a 20-item self-report measure designed to assess paranoid ideation in response to everyday events and situations. Items on the PS include 'someone has it in for me' and 'it is safer to trust no one'. Participants rate their agreement with each item using a 5-point scale (1 = *not at all applicable to me* to 5 = e*xtremely applicable to me*). Individual scores are summed, ranging from 20 to 100 with higher scores indicating greater endorsement of paranoid ideation. The PS has been used widely within clinical research and has good psychometric properties with internal consistency, $r = 0.84$, and test–retest reliability, $r = 0.70$ [33]. The LEAP contributed to discussion and decision-making around the choice of primary outcome, informed by the data collected from the feasibility study [22].

**Table 1** Definitions of adherence

|  | Definition | Acceptable values | Target values |
|---|---|---|---|
| **Individual sessions** | To 'complete' an individual session participant must… | Enter responses on at least 20 out of 40 trials (50%) | Enter responses on at least 30 out of 40 trials (75%) |
| **Intervention overall** | To 'complete' the intervention participants must… | Complete ¼ of all sessions scheduled to date | Complete ½ of all sessions scheduled to date |
| **Number of participants** | Proportion of sample to be adherent to meet overall trial adherence targets | 50% | 75% |

**Table 2** Trial assessment schedule

| | Timeline | Screening | Baseline 0 weeks | Sessions 1-6 | Follow-up 6 weeks (this is end of treatment for the 6-session arm, mid treatment for the other arms) | Sessions 7-12 (The 6-session arm does not receive these treatment sessions, but does receive assessments) | Follow-up 12 weeks (this is 3 mth post randomisation FU for 6 session arm, end of treatment for others) | Follow ups 18 weeks and 24 weeks |
|---|---|---|---|---|---|---|---|---|
| 1 | *Socio-demographics* | | | | | | | |
| 2 | Inclusion/Exclusion criteria checklist | ■ | | | | | | |
| 3 | Capacity Assessment Tool | ■ | | | | | | |
| 4 | Socio-demographic data | | ■ | | | | | |
| 5 | Participant Consent Form | ■ | | | | | | |
| 6 | Participant Information Sheet | ■ | | | | | | |
| 7 | Credibility/Expectancy Questionnaire | | ■ | | | | | |
| 8 | Wechsler Test of Adult Reading (WTAR) | | ■ | | | | | |
| 9 | Persuadability subscale (PER) | | ■ | | | | | |
| 10 | *Target mechanism (bias)* | | | | | | | |
| 11 | Similarity Ratings Task (SRT) | ■ | C/bal ↑ | ■ ← | ■ | ■ | ■ → | C/bal ↑ |
| 12 | Scrambled Sentences Task (SST) | | C/bal | ■ | ■ | ■ | ■ | C/bal |
| 13 | *Clinical Symptoms* | | | | | | | |
| 14 | Paranoia Scale* | | C/bal | ■ | ■ | | | C/bal |
| 15 | Green Paranoid Thought Scale (GPTS) short version (18 items) | | C/bal | ■ | ■ | | | C/bal |
| 16 | PANSS- item 6 only | ■ | ■ | | | | ■ | |
| 17 | SCID-5 Diagnostic interview | | ■ | | | | | |
| 18 | HADS | | ■ | | ■ | | ■ | ■ |
| 19 | Paranoia Worries Questionnaire | | ■ | | ■ | | ■ | ■ |
| 20 | *State Mood change* | | | | | | | |
| 21 | VAS (anxious, sad, paranoid, friendly) | | | Pre-post session | | Pre-post session | | |
| 22 | *Recovery* | | | | | | | |
| 23 | WEMWBMS | | ■ | | ■ | | ■ | ■ |
| 24 | QPR | | ■ | | ■ | | ■ | ■ |
| 25 | ReQoL | | ■ | | ■ | | ■ | ■ |
| 26 | User Satisfaction (Ben-Zeev et al 2014) | | | | ■ | | ■ | FU 24w only |
| 27 | *Miscellaneous* | | | | | | | |
| 28 | Treatment As Usual | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| 29 | Self-reported Treatment Allocation | | | | | | | FU 24w only |
| 30 | Negative Effects Questionnaire | | | | ■ | | ■ | |

1. Grey cells indicate measures subject to counterbalancing (C/bal) for:

    a) task order (indicated by vertical arrows)

    b) item content across multiple uses of the same task (indicated by horizontal arrows)

2. Full counterbalancing schedules available upon request from the corresponding author

3. Pre-post session = this assessment is given before (pre) and after (post) every session

4. FU 24w only = this assessment is administered at 24 week follow up only

**Secondary outcome measures**

The LEAP contributed to discussion and decision-making around the choice of secondary outcomes informed by the data collected from the feasibility study [22].

1. *Paranoid symptoms* are also measured as secondary outcomes at baseline, 6, 12, 18 and 24 weeks post-randomisation (24-week measurement is the primary outcome).

2. *Paranoid beliefs* (i.e. interpretation bias, the cognitive mechanism targeted by the intervention) are measured at baseline, 6, 12, 18 and 24 weeks post-randomisation and after every session using:

    • *Similarity Rating Test* (SRT)
    • *Scrambled Sentences Task* (SST)

3. *Clinical symptoms* are measured at baseline, 6, 12, 18 and 24 weeks post-randomisation using:

    • *Green Paranoid Thoughts Scale Revised* (R-GPTS).
    • *Positive and Negative Symptom Schedule* (PANSS)—item 6 only
    • *Hospital Anxiety and Depression Scale* (HADS)
    • *Paranoia Worries Questionnaire* (PWQ)

4. *State mood changes* are measured before and after each session using visual analogue scales (VAS: anxious, sad, paranoid, friendly). VAS for mood are psychometric response scales used to measure present mood state using visual representations (e.g. faces representing different mood states) [34]. These detect any abrupt changes in mood as part of the safety assessment protocol of the STOP mobile app.

5. *Recovery* is measured at baseline, 6, 12, 18 and 24 weeks post-randomisation using:

    • *The Warwick-Edinburgh Mental Wellbeing Scale* (WEMWBS)

- *Questionnaire about the Process of Recovery* (QPR)
- *Recovering Quality of Life* (ReQoL)
- *User satisfaction* with the mobile app is measured with the Ben-Zeev et al. [35] instrument.

6. Miscellaneous

- *Negative Effects Questionnaire* (NEQ)
- *Adverse events* (AEs) data is also collected and reported using an adverse events checklist that was developed specifically to assess the safety of STOP; details about the development and content of this checklist have been published elsewhere [36].

See Supplementary file 1 for further details on the scales used for the secondary outcomes.

**Socio-demographic data and other baseline measures**

The following socio-demographic data are collected at baseline: age, sex at birth, ethnicity, the Wechsler Test of Adult Reading (WTAR) [37] as a proxy measure of IQ, educational level, employment, living arrangement, relationship status, self-reported dyslexia and age of onset of distressing paranoia. The subscale of the Multidimensional Iowa Suggestibility Scale [38] is administered at baseline as part of the safety protocol to determine those who require weekly, as opposed to fortnightly, check-in calls (see Sect. 13, 'Participant timeline'). Two items ('By the end of the intervention, how much improvement in your paranoid symptoms do you think will occur?' and 'By the end of the study, how much improvement in your symptoms do you really feel will occur?') from the Credibility and Expectancy questionnaire [39] are used to measure participants' expectations of benefiting from taking part in the trial,
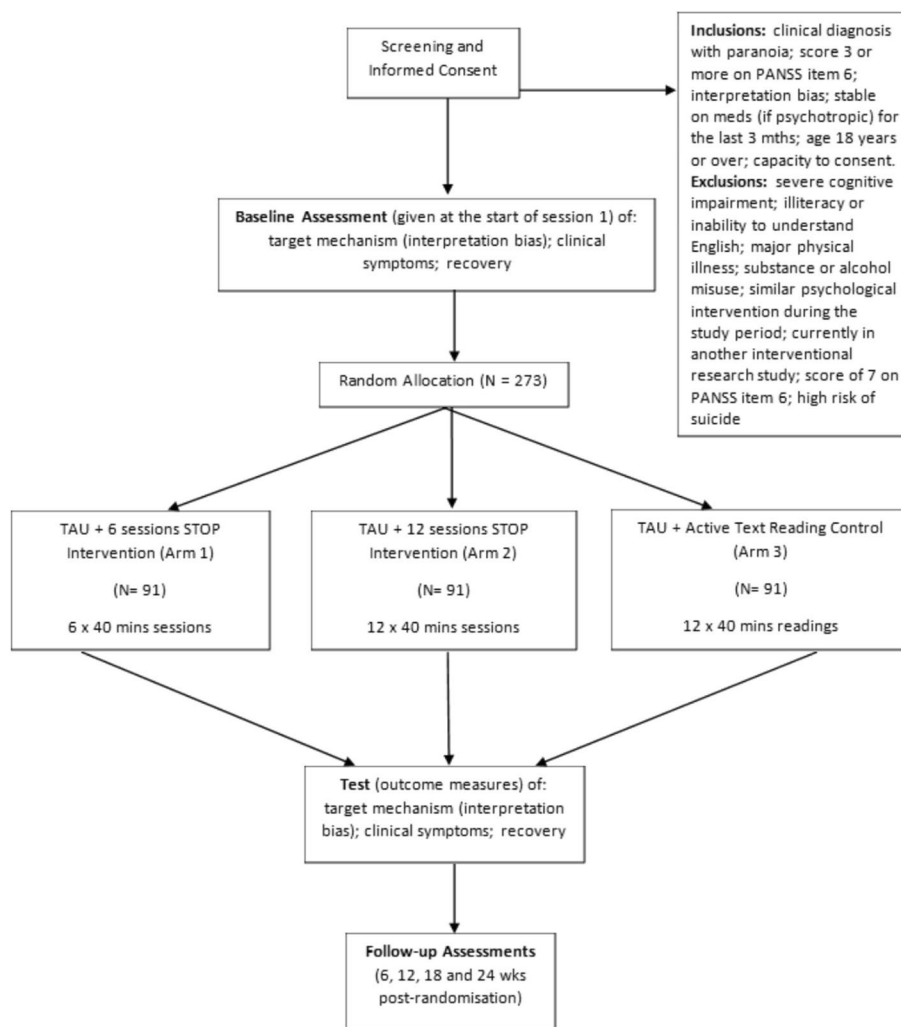


**Fig. 3** Trial flow diagram

to assess both equipoise and the role of expectation in predicting improvement.

### Participant timeline {13}

The participant pathway through the trial is shown in the trial flow diagram (Fig. 3). Numbers indicate recruitment targets.

Participants interested in taking part are provided with the participant information sheet (PIS) and given at least 24 h to consider taking part. They also have the opportunity to ask any questions before written or electronic consent (as the participant prefers) is obtained. Informed consent is taken before screening due to the significant and necessary data collection that screening involves. Participants' eligibility is established by systematically assessing the inclusion and exclusion criteria described above. Eligible individuals who choose to take part then meet remotely with a trained researcher who administers the baseline assessment. Researchers then randomise the participant using the blinded randomisation service provided by King's Clinical Trials Unit (KCTU). Researchers have remote contact with participants to carry out the follow-up assessments.

### Check-in calls

All participants (intervention and control) receive a fortnightly phone call from researchers, following a pre-prepared script to limit bias. Participants who score > 58 (2.5 standard deviations above the general population mean) on the persuadability subscale of the Multidimensional Iowa Suggestibility Scale (MISS) [38] are deemed more vulnerable to the possible negative effects of the intervention by the Medicines and Healthcare products Regulatory Authority (MHRA) and are therefore offered weekly check-in phone calls. Key objectives of the check-in calls are (i) to check participants well-being, (ii) record any adverse events and (iii) promote engagement in the trial.

### Sample size {14}

With alpha = 0.05 and 80% power, recruiting a sample of 77 per group would permit the detection of a clinically useful drop of 8 points on the paranoia scale compared to TAU assuming a standard deviation of 17.5 (based on feasibility data) [22]. This magnitude drop would be a 15% reduction in the average score of paranoid patients [40, 41] and represents an effect size of $d = 0.46$ (small to medium). Assuming a conservative estimate of 15% drop-out (from feasibility data) a total sample size of 273 is required when using an independent $t$-test.

### Recruitment {15}

Participants include individuals who are, and are not, currently using mental health services. The trial therefore has a wide range of recruitment pathways including:

- NHS Trust sites: 20 NHS Trusts across England and Wales contribute to the study through the CRN portfolio. In these trusts, GCP-trained CRN and clinical staff may refer individuals to the central team for screening or, following suitable training from the central team, may perform screening themselves and refer on eligible participants.
- McPin Foundation: dedicated staff members from McPin contribute significantly to recruitment through service user networks, attending nationwide events and other study publicity.
- Recruitment registers in secondary care comprising individuals who have agreed in principle and consented in advance to a direct approach from research teams.
- MQ Mental Health Research (https://www.mqmentalhealth.org/home/). A UK charity which publicises opportunities to get involved in research directly to the general public.
- *Self-referral* through the study website could result from any of the study publicity activities.

Potential participants from any source are encouraged to access preliminary or additional information about the study through the study website and recruitment flyers.

## Assignment of interventions: allocation
### Sequence generation {16a}

Randomisation is at the individual level in the ratio 1:1:1 and is performed by software development company, Avegen, with advice and input from King's Clinical Trials Unit (KCTU). The randomisation procedure is completed in the STOP app. There are no adjustments permitted to randomisation once completed. Researchers and participants remain masked to the allocated condition. Block randomisation with randomly varying block sizes is used.

*Randomisation stratifiers* (measured at screening) are as follows:

1. Study hub (Bath, London)
2. Sex at birth (male, female)

### Concealment mechanism {16b}

Randomisation lists including participant PIN numbers are generated and communicated by KCTU to the software developer. The developer is responsible for implementing the randomisation process within their software

platform. The patient's assignment to study arm remains concealed from all those involved in the study except two nominated members of the software development company. Data output for the data monitoring committee (DMC) and trial steering committee (TSC) is pooled across arms. Upon specific request of the DMC, partially unblinded data can be generated with labels A, B and C to denote the different arms and only the external committee members and partially unblinded trial statistician see these data. The study team, including the senior statistician, sees only collapsed data.

### Implementation {16c}
KCTU provides a unique blocked set of randomisation codes to the developer. Prior to randomisation, each participant is registered on a database system customised for the study (MACRO) which provides a unique participant identification number (PIN) that is used for their registration into the STOP app. PIN allocation is undertaken by trial researchers going to www.ctu.co.uk and clicking the link to access the MACRO system. A full audit trail of data entry is automatically date and time stamped, alongside information about the user making the entry within the system.

## Assignment of interventions: blinding
### Who is blinded {17a}
This is a double-blinded study meaning that both researchers and trial participants are unaware of which arm a participant has been randomly allocated to. Two members of the developer team who are independent of those working with the trial research team are designated to be fully unblinded throughout the trial and these individuals are not permitted to have direct contact with any members of the research team.

Any cases of inadvertent unblinding (e.g. researcher becoming aware of the content of sessions during a routine phone call with a participant) are recorded by researchers as protocol violations and participants continue with their participation. During data analysis, sensitivity analyses are performed to assess the impact of unblinded cases.

### Procedure for unblinding if needed {17b}
It is not anticipated that there would be a need for an emergency code break during the trial. However, if there is a compelling reason to break the blind (e.g. in the event of a serious adverse event which is related to the study intervention, procedures or device) on an individual basis, then this would be decided on the basis of advice from the DMC and implemented via an approach from the CI to one of the developer's two named randomisation list holders.

## Data collection and management
### Plans for assessment and collection of outcomes {18a}
Study researchers at both hubs are provided with appropriate training on the protocol for collecting trial data. See the 'Trial design' section {12} for an overview of the assessment schedule.

The design includes every arm receiving weekly post-session 'interim' assessments for 12 weeks. Interim assessments comprise a small subset of the main outcome measures which are completed unsupervised within the STOP app, as follows:

- Paranoid beliefs measured using an 8-item version of the SRT and the SST [26].
- Clinical symptoms assessed by Paranoia Scale [33] and R-GPTS [42].
- State mood change measured pre-post each session using visual analogue scales (VAS; anxious, sad, paranoid, friendly) [34].

Participants also receive a check-in call on a fortnightly basis (weekly if they are high risk) to complete the adverse events checklist. The rationale for and development of this checklist is published elsewhere [43].

Follow-up assessments are given at 6, 12, 18 and 24 weeks post-randomisation to measure longer-term effectiveness of the intervention. Follow-up assessment includes the full battery of outcome measures at 6, 12, 18 and 24 weeks as shown in Table 2.

### Plans to promote participant retention and complete follow-up {18b}
A bespoke schedule of email and mobile app reminders to both researchers and participants is programmed as part of STOP, as detailed in Table 3.

Participants receive £20 for each follow-up assessment they complete (four in total) as reimbursement for their time. They receive an additional £20 final payment upon completion of the last assessment.

Participants who withdraw early from the study may have consented, via a dedicated checkbox, to an invitation to complete a short exit assessment. In this case, participants who remain willing to do so at the time of withdrawal will:

- Self-report the main reason for withdrawal, which will be qualitatively recorded by a researcher in the withdrawal care report form (CRF) booklet
- Complete Ben-Zeev et al. [35] mobile app user satisfaction questionnaire
- Complete a bespoke, brief five-question questionnaire created for the study

**Table 3** Schedule of reminders and notifications

| Reminder/notification  (sent 9am on next calendar day) | To participant | | To researcher |
|---|---|---|---|
| | Push notification (in app) | Email | Email and researcher platform alert |
| Day before a scheduled session | x | x | |
| Incomplete session:<br>• Session not started on the scheduled day<br>• Session only partially completed on the scheduled day | x | x | x |
| Missed session: participant omits or only partially completes scheduled session after the scheduled day | x | x | x |
| Mood worsening across 3 consecutive sessions on any VAS scale | | | x |
| Day before a scheduled follow-up (6, 12, 18, 24 weeks) | x | x | x |

Beyond this, no further data are collected for any participants who withdraw from the study and no follow-up data are collected.

In contrast, for those who deviate from the protocol, or withdraw from the study intervention only, all follow-up data are sought.

### Data management {19}

A large proportion of study data is collected on the STOP app or by direct entry into the software company's researcher platform (HealthMachine: HM™). Monthly data outputs of raw, non-aggregated data points are provided to the trial's statistician by the developer. Full data extracts are also sent every 6 months to prepare a comprehensive report for the DMC in order to meet DMC reporting requirements or whenever requested by DMC. In addition, data related to study progress (e.g. eligibility, recruitment, retention, drop-out, etc.) and study milestones (e.g. adherence) are available live via either a live dashboard linked to the app and/or internal study recruitment tracking records.

Data not captured by the software company's systems are managed using KCTU procedures. A bespoke web-based electronic data capture (EDC) system has been designed, using the InferMed Macro 4 system. Database access is strictly restricted through user-specific passwords to the authorised research team members. Participant initials and age are entered on the EDC. No data is entered onto the EDC system unless a participant has signed a consent form to participate in the trial. The CI team undertakes reviews of the entered data, in consultation with the project analyst, for the purpose of data cleaning and request amendments as required. Upon a participant's exit from the trial, a data monitor (CI or project manager) reviews all the data for each participant and provides electronic sign-off to verify that all the data are complete and correct. When all participants have received sign-off, data are formally locked for analysis and KCTU provides a copy of the final exported dataset in.csv format to the CI for onward distribution.

### Confidentiality {27}

All information provided by participants is kept confidential, within the limits of the law (i.e. the Data Protection Act). The limits of confidentiality and possibility of action arising from certain disclosures (e.g. participant discloses information that puts themselves or someone else at risk of serious harm) are outlined in the PIS and comprise part of the informed consent process.

Throughout the active trial period, all data collected are held in pseudonymised form on secure password-protected university servers. This includes indirect identifiers (e.g. age, gender, ethnicity and sensitive clinical diagnostic information). Individual participants are coded within the datasets using their PIN assigned at study registration. Direct identifiers (e.g. name, contact details, etc.) are stored in a separate file location on the university servers and used only for the purpose of liaising with participants/clinicians during participation in the trial. Only fully anonymised aggregated quantitative data are used when the results of the study are published or otherwise disseminated so participant identity will never be revealed.

The quantitative data generated by this study are suitable for sharing after having been fully anonymised and are deposited with UK Data Archive. Preparation for this purpose is informed by recent literature advising on the best way to reduce the possibility of re-identification [41, 42, 44]. We include a specific item in our informed consent procedures allowing participants to indicate whether they consent to anonymous data sharing via a national repository. The data deposit is limited to those cases giving consent.

**Plans for collection, laboratory evaluation and storage of biological specimens for genetic or molecular analysis in this trial/future use {33}**

Not applicable as there will be no biological specimens collected (see above 26b).

## Statistical methods

### Statistical methods for primary and secondary outcomes {20a}

*Trial objective*

We are interested in the estimand for the effect of treatment under a treatment policy initially assigned at baseline as expected under routine clinical care—as defined by International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) [44]—at prespecified follow-up times. Following the ITT (intention to treat) principle (which corresponds to the treatment policy strategy approach), this requires estimating the ITT estimand, which is defined as the average treatment effect (reduction of paranoid symptoms measured using the paranoia scale) between 6 and 12 sessions versus control at the 24-week follow-up, regardless of adherence to the allocated study treatment (or other intercurrent events that could occur) for all randomised individuals. This has the consequence that subjects allocated to a treatment arm need to be followed up, assessed and analysed as members of that arm irrespective of their compliance to the planned course of treatment [44]. This approach requires the collection of outcome data after treatment discontinuation, and we are required to treat missing data after treatment discontinuation as though they had been observed [44].

*Baseline comparability of randomised groups*

Baseline variables will be described overall and by intervention group. Frequencies and proportions will describe categorical variables, while numerical variables will be described using mean and standard deviation (SD) if normally distributed and median and interquartile range (IQR) if not normally distributed.

No statistical testing of the baseline differences between randomised groups is done.

*Descriptive statistics for outcome measures*

The primary and secondary outcomes will be summarised at baseline, 6 weeks, 12 weeks and 24 weeks post-randomisation by intervention group and overall. Mean, SD, medians and IQR will be used to summarise variables, respectively.

*Inferential analysis*

The main statistical analyses aim to estimate the group mean differences and associated 95% confidence intervals

at the 6th week (end of the 6-session treatment), 12th week (end of 12 session treatment), the 18th week and the 24th week post-randomisation (6 weeks and 12 weeks post-treatment observation time). These are estimated using models that account for the repeated measurements of the participants, as described below.

Randomisation is stratified by hub (London and Bath) and sex at birth, which will always be included as covariates in all statistical models. Post-randomisation measurements will be taken at 6, 12, 18 and 24 weeks. We expect an increasing dropout rate as treatment progresses, and as a result, we will analyse all four time points simultaneously, under (restricted) maximum likelihood, to reduce potential biases and to maximise power.

*Analysis of primary outcome*

The primary outcome is mean paranoid symptoms measured using the Paranoia Scale (PS) [33] at 24 weeks post-randomisation. PS ranges from 20 to 100 and are treated as continuous. Our feasibility study suggested that the primary outcome is normally distributed and ranges between 20 and 80 points.

To evaluate the effect of a 6- and a 12-session STOP mobile app compared with a 12-session text reading control, a linear mixed-model repeated measures (mmrm) approach [41] will be fitted using restricted maximum likelihood (reml). The model will allow a separate mean paranoia scale (PS) parameter at each assessment time in each treatment group and an unstructured covariance matrix of the response-level residuals.

Under the repeated measures setup, we will treat baseline outcome and dummy variables for the randomisation factors (study hub, sex at birth) as independent variables with a design matrix that allows no treatment difference at baseline (though a different residual variance at follow-up with non-zero covariances over time and across treatment groups). Therefore, treatment randomisation group (with levels denoting control, 6-session and 12-session treatment), time (with levels denoting weeks 6, 12, 18 and 24 weeks), baseline PS score, treatment by time interaction, baseline PS score by time interaction, study hub and sex at birth will constitute the fixed part of the model. Continuous baseline covariates will be centred at their means. This setup allows us to estimate the treatment effect at 24 weeks post-randomisation and lets follow-up data be treated as ignorable under a missing-at-random assumption. At the same time, the model accounts for the possible imbalance due to random sampling in baseline measurement of the outcome variable to control for pre-treatment differences.

In addition to mean treatment differences with 95% confidence intervals, standardised effect sizes (Cohen's *d* calculated as estimated treatment difference divided

by the standard deviation of baseline outcome) with 95% confidence intervals will be presented.

### Model assumption checks

The linear mixed-effects models assume normally distributed residuals. Outcome residuals for each time point will be plotted using a Q–Q plot to check for normality and the existence of outliers. If violations of the assumptions will be observed, bootstrapped 95% confidence intervals will be calculated using the bootstrap procedure for multilevel data.

### Analysis of secondary outcomes

For the secondary outcomes, we will use the same model to evaluate the effect of the STOP intervention, provided that the outcome values can be assumed to be normally distributed. Where this is not the case, we will make use of generalised random effects models with robust standard errors, under the missing-at-random assumption about the data. For binary outcomes, we will fit random effects logistic models and for ordered outcomes random effects ordinal logistic models command.

### Method for handling multiple comparisons

We control for type 1 error for our two planned and orthogonal contrasts (6 months against control and 12 months against control at 6 weeks follow-up) by using LSD test procedure, which does not require a correction of the familywise error of pairwise comparisons of three groups if the main effect is significant. No correction will be made for multiple comparisons involving the primary outcome.

Apart from our main outcome (paranoia scale), we have identified five key secondary outcomes (two measures of the target mechanism: SRT, SST; three measures of clinical symptoms: GPTS, HADS, PWQ) upon which we will perform hypothesis testing.

For the remaining variables, we will provide statistics and estimated treatment effects as purely explorative results and do not perform formal statistical tests. Standard errors will be provided for the reader to make their own formal assessment if desired.

### Analysis of safety

There is a small likelihood that the STOP device can trigger paranoia symptoms in users over and above their baseline incidence levels of paranoia. We will therefore analyse AEs/SAEs which involve paranoia symptoms separately from other AEs/SAEs to identify potential increases.

We will assess whether each AE and SAE is related to the app itself and present any AEs and SAEs classified as potentially related to the app (including all paranoia symptoms) in the following ways:

- For SAEs, each event, date and reason for potential relation to the app will be listed individually.
- We will then (i) plot the number of adverse events and SAEs per predefined time period (e.g. $6 \times 4$ week blocks) and (ii) produce a cumulative plot over time. The same descriptive analyses will be presented per arm. The events will be presented as collapsed AE and SAE as well as grouped into main categories. The individual events and numbers will be regularly assessed by the CI and by independent members of the DMC at each meeting. The DMC will also assess any critical development between bi-annual meetings when necessary.

We do not expect the number of potential app-related AEs/SAEs to occur frequently enough to perform formal statistical analyses during the trial and will only report the incidence rate (total number of those having the event divided by the person-months at risk) and the ratio of incidence rates between the two treatment arms per time period will be calculated to allow detection of any safety concerns within the treatment arm.

The data from the adverse events checklist in the 12-session arms (intervention and control) will be analysed using qualitative content analysis to produce a list of adverse events along with their frequencies, seriousness, relatedness to the app and possible methods of prevention/mitigation. Additionally, we will examine the demographic and clinical characteristics of those who experienced adverse events to identify patients who might be at a higher risk.

The NEQ data will be provided as a total/overall score of the negative effects that are related to treatment and those that are related to other circumstances. As these data provide quantitative scores for every participant, we will be able to perform comparative statistical tests between treatment and control arms using a linear regression approach.

### Sensitivity of results to potential imbalance of baseline characteristics

To assess the effect of potential chance imbalance five prespecified baseline variables (education level, age, ethnicity, paranoia severity and bias score (SRT)) will be controlled for and included as additional covariates in the models. Any changes in treatment differences will be reported.

## Interim analyses {21b}

No interim analyses are planned.

## Methods for additional analyses (e.g. subgroup analyses) {20b}

No subgroup analyses are planned.

## Methods in analysis to handle protocol non-adherence and any statistical methods to handle missing data {20c}

### Sensitivity analysis due to protocol violations

A protocol violation or non-compliance is any deviation from the protocol. These are recorded in a dedicated log.

The statistical analyses will be primarily concerned with two protocol violations:

1) Violations with regard to the occurrence of a session and the extent of the adherence to the study, as described in Section {11c} 'Strategies to improve Adherence to Interventions'
2) Violations due to unblinding

We will assess protocol violations by assessing the violations (e.g. unblinding or data collected outside the collection time frame) in further sensitivity analysis by supplementing the model of the primary analysis with an additional binary indicator variable coded 1 for the existence of any protocol violation and 0 otherwise. The analysis of this model will give intervention effect estimates adjusted for potential effects of protocol violation. We will report the changes in the predicted outcome differences alongside the main analysis.

### Reporting of missingness

We will follow CONSORT guidelines of reporting the reasons why patients were lost to follow-up and hence why outcomes are missing. To identify the potential for bias due to missing data, we will present a table comparing the distribution of baseline data by treatment arm for patients with observed follow-up data and for patients with missing follow-up data separately for each study arm.

Missing data will be described both as a proportion in each variable separately as well as using a graph that ranks variables from most complete to least complete. This will be done for all participants together as well as for each arm separately.

### Missing data in baseline variables

If the proportion of missing data in baseline variables is small (approximately 1–2%), we will impute in with the mean of the baseline variable [45]. If the proportion of missing data in the baseline variables is large (>2%), we will consider using multiple imputations [46] by chained equations, which can accommodate the imputation of missing data in baseline variables as well as in outcome measures at the same time. Results from this process will be contrasted with results from the complete case data analyses.

### Missing data in the outcome

The described mixed model will be estimated via restricted maximum likelihood that is valid under the missing at random (MAR) assumption and allows to include all patients with at least one follow-up observation. The MAR assumption requires that all predictors of the missingness mechanism be included in the models to maximise the likelihood that, conditional on these predictors, the outcome is missing at random.

To make the missing at-random assumption more credible, we will identify prespecified baseline variables (age, sex at birth, ethnicity, IQ (WTAR), educational level, employment, living arrangement, relationship status, self-reported dyslexia, age of onset of distressing paranoia, paranoid beliefs, paranoia scale and HADS) as potential predictors of missing data. We will perform logistic regressions to identify significant predictors of missing outcome variables, following Carpenter and Smuk's guidelines [47]. The analysis incorporates constructing binary indicators for missing outcomes at various time points and utilizes logistic regression models to assess if baseline variables predict missing outcome data. The statistically significant variables at a 10% level from this will constitute the predictors of missingness and will be incorporated in the mixed-model repeated measures models for the analysis of the primary or secondary outcomes as part of a sensitivity analysis. Interactions between covariates and time will be assessed relaxing the assumptions that missingness patterns are similar between the treatment arm and over time. We will report how sensitive our results are to the inclusion of these variables and post-treatment group difference estimates and associated confidence intervals will be reported.

### Sensitivity of results due to missing data by assuming missing not at random (primary analysis)

Analysis of data where the outcome is incomplete always requires untestable assumptions about the missing data, commonly that the data are missing at random. The estimand would be inappropriate if a large proportion of participants leave the study due to intercurrent events and are not available for the final endpoint measurement. In this case, we assume that the data are potentially not missing at random (MNAR). To address this, the primary analyses are conducted using a pattern mixture with multiple imputations model [48]. This approach formulates the MNAR problem in terms of the different distributions between the missing and observed data

and different distributions can be specified for different patterns of missingness of (groups of) participants that reflects a specific assumption appropriate to their treatment arm, drop-out time and possibly other relevant information.

We will perform sensitivity analyses to explore the effect of departures (varied over a plausible range) from the assumption of missing at random made in the main analysis following the steps recommended by White et al. [49]. We will set up a multiple imputation model consistent with the primary substantive analysis model. We will use multiple imputations to impute missing values under a MAR assumption and modify the MAR imputed data to reflect a plausible range of patterns related to the effect of non-compliance followed by drop-out following the guidelines of Carpenter et al. [48].

### Missing items in scales and subscales
The planned strategy for handling missing data at the item and scales level will depend on the amount of missing data observed and the planned analyses for the outcomes. For missing item data in secondary outcomes and baseline measures, we will use pro-rating if less than 20% of items are missing [50, 51]. If more than 20% are missing, we will treat it as a missing data point. To ensure the same strategy is followed across all scales reported in the principal paper(s), any guidance given by authors of validated questionnaires will supersede the methods outlined herein.

### Plans to give access to the full protocol, participant level-data and statistical code {31c}
The full statistical analysis plan will be published in an open access forum or journal. Data files are to be deposited at the UK Data Archive. This is a clinical trial registry recognised by WHO and ICMJE that accepts planned, ongoing or completed studies of any design. It provides content validation and curation and the unique identification number necessary for publication. All study records in the database are freely accessible and searchable. The statistical code is to be deposited in GitHub (www.github.com), a Git repository hosting service.

## Oversight and monitoring
### Composition of the coordinating centre and trial steering committee {5d}
The design, development and implementation of the trial is managed by the CI and the co-applicants, in consultation with service users. A dedicated project manager, reporting to the CI, oversees trial conduct, assists with governance and regulatory matters and provides quality control. A Project Management Group (PMG) meets quarterly to ensure the project remains focused on

delivering against its objectives and monitors the project against its stated budget, schedule, milestones and risks. Its membership includes the investigators, the project manager, two trial statisticians, a member of the software development company and a representative from The McPin Foundation. It will oversee the preparation of reports to the funder, TSC and DMC.

The purpose of the TSC is to oversee the trial on behalf of the sponsor and funder; to monitor adherence to the protocol and statistical analysis plan; to ensure the rights, safety and well-being of the participants and to provide independent advice on all aspects of the trial. It consists of five members independent of the trial team (a chairperson, a statistician, two clinical academics, and a patient and public involvement (PPI) representative) and five members of the trial team (CI, two trial statisticians, project manager and an additional PPI representative).

### Composition of the data monitoring committee, its role and reporting structure {21a}
The DMC consists of three independent members: a chairperson, a clinical expert, and a statistician. The DMC monitors the data emerging from the trial as they relate to data quality and the safety of participants and may request partially unblinded, fully unblinded or interim analyses of trial data to be undertaken. It makes recommendations to the TSC on whether the trial should continue, stop or be modified.

### Lived experience advisory panel (LEAP): context, formation and contribution
A STOP study lived experience advisory panel (LEAP) was recruited in March and April 2021. The LEAP consists of five active members (three men and two women) who use their lived experience of paranoia in consultation with the STOP study team. The LEAP is coordinated by McPin staff who also use their lived experience of paranoia in the project work. It meets four times a year (every 3 months), mostly online but also in person.

The STOP predecessor study, cognitive bias modification for paranoia (CBM-pa), also had a LEAP. The CBM-pa LEAP developed early ideas for the STOP study, including branding guidance to be used on all study-related materials such as the website, newsletter, and within the app itself.

One of the first ways of involving the newly formed STOP LEAP was to have them help develop and review the scenarios used in the STOP app. The LEAP also participated in user testing sessions where their ideas about the use and content of a prototype of the app were received by the study team and researchers, including new features such as gamification. The STOP LEAP contributes to study newsletters and was consulted about

how much they thought participants should be paid for taking part. The group was consulted regarding the outcome measures in use in the study trial.

The LEAP is key in helping with recruitment to the trial. LEAP members use their own local networks and knowledge of community organizations to promote and recruit participants to the STOP trial using a study poster and flyer. The LEAP will be asked to contribute to the interpretation of the trial outcome data and for ideas for dissemination of the study findings.

### Adverse event reporting and harms {22}

Full details on the process followed and materials developed for identifying, recording and reporting adverse events in this trial have been published elsewhere [43].

#### *AE data collection*

Information concerning AEs is collected using an adverse events checklist. Researchers use this checklist to proactively enquire about patient safety at every scheduled contact (i.e. fortnightly check-in calls and follow-up assessments).

#### *Assessing AE data*

Every AE reported by participants and captured on the checklist is assessed for seriousness, relatedness, severity and whether it was anticipated or not, by trial researchers.

#### *Reporting AE data*

The study follows European Commission guidance on serious adverse event reporting in clinical investigations [52] and by the HRA [53]. The trial also follows its locally developed procedures for identifying, recording and reporting Adverse Events [43].

#### *Regular reporting*

All AEs and SAEs are pooled and reported to each meeting of the DMC, who report to the TSC. As this is a medical device trial, quarterly safety summary reports detailing all SAEs occurring in the last quarter must be submitted to MHRA. Every SAE must also be individually notified to MHRA using their standard reporting template, within 48 h of it coming to the attention of the CI. SAEs are also reported to the sponsor who instructs the team whether or not to notify the Research Ethics Committee (REC).

### Frequency and plans for auditing trial conduct {23}

The TSC and the DMC are responsible for ensuring the trial is conducted to a high standard. These committees convene every 6 months. The trial statistician prepares a comprehensive report for each DMC meeting.

### Plans for communicating important protocol amendments to relevant parties (e.g. trial participants, ethical committees) {25}

We do not envisage any amendments to the protocol due to the implications on time and resources. Any amendment would be submitted to the TSC for approval in principle and subsequently to REC and MHRA for governance approval. Subsequent changes would then be communicated to sites and, if appropriate, participants and recorded in the ISRTCN registration.

### Dissemination plans {31a}

Results will be published in good quality academic journals and presented at appropriate national and international conferences. The study website (www.STOPtrial.co.uk) will be used to communicate study progress and key findings. The study and its findings will also be publicised through the dissemination networks at The McPin Foundation (https://mcpin.org/), using materials such as infographics, leaflets, podcasts and articles for newsletters and popular press. These are made available on the study website and are proactively disseminated through user networks, relevant mental health charity websites and newsletters, and NHS trust communications.

### Discussion

The development of digital mental health interventions (DMHIs) has seen exponential growth over the last 20 years [54]. Despite some growing evidence around their effectiveness [55], some reviews have highlighted that only a few of these interventions are evidence-based [56]. For that, NICE (2023) has made recommendations regarding the use of eight digital therapies for managing mental health conditions like agoraphobia, anxiety, depression and psychosis [57]. For DMHIs to be integrated into health services and realize their benefits for patients, it is crucial to strengthen the evidence base demonstrating their efficacy. Additionally, assessing safety in digital therapeutic trials has become paramount now that they are uniformly classified as medical devices. The current study will build on a previous feasibility study [22] by evaluating the efficacy and safety of a novel, entirely self-administered digital therapeutic for paranoia called STOP. If effective, STOP could complement other psychological treatments (including CBT) [58] offering an accessible and flexible supplementary treatment.

### Trial status

Protocol version 3.2, 12 September 2022. Recruitment for the trial commenced in September 2022. Recruitment ended in June 2024 when the recruitment target was met.

## Abbreviations

| | |
|---|---|
| AE | Adverse event |
| SAE | Serious adverse event |
| CBM-pa | Cognitive bias modification in paranoia |
| CBT | Cognitive behaviour therapy |
| CI | Chief investigator |
| CRN | Clinical research network |
| DMC | Data monitoring committee |
| GCP | Good Clinical Practice |
| LEAP | Lived experience advisory panel |
| MHRA | Medicines and Healthcare products Regulatory Authority |
| MRC | Medical Research Council |
| NHS | National Health Service |
| NICE | National Institute for Health and Care Excellence |
| PMG | Project Management Group |
| PIN | Participant identification number |
| PPI | Patient and public involvement |
| REC | Research Ethics Committee |
| RCT | Randomised controlled trial |
| STOP | Successful treatment of paranoia |
| TAU | Treatment as usual |
| TSC | Trial steering committee |
| WTAR | Wechsler Test of Adult Reading |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13063-024-08570-3.

> Supplementary Material 1.  [ 25–27, 34–37, 42, 59–69 ].

## Authors' contributions {31b}

JY is the chief investigator of the study, conceptualised and implemented the study design, secured study funding, coordinated the writing of the study protocol, contributed to trial management, supervision, reviewing and editing the manuscript. CH: recruitment and testing of trial participants, data collection, scoring and cleaning, contributed to reviewing and editing the manuscript. CF: recruitment and testing of trial participants, data collection, scoring and cleaning, contributed to reviewing and editing the manuscript. CWH: contributed to the study design, supervision, reviewing and editing the manuscript. DS: is the lead statistician, carried out the power calculation, prepared the statistical analysis plan and assisted with study design decisions. EM: contributed to the study design, supervision, reviewing and editing the manuscript. EP: contributed to the study design, supervision, reviewing and editing the manuscript. GV: is a trial statistician, contributed to the statistical analysis plan and is responsible for the implementation of analyses. PM: contributed to the study design, reviewing and editing of the manuscript. PJ: principal investigator at Bath, contributed to the study design, supervision, reviewing and editing of the manuscript, and trial management at Bath. RT: recruitment and testing of trial participants, data collection, scoring and cleaning and led the preparation of the manuscript's first draft. TK: contributed to patient and public involvement, study design, supervision, reviewing and editing of the manuscript. TR: contributed in reviewing and editing the manuscript, project management and set-up, and recruitment. SS: contributed to the study design, supervision, reviewing and editing of the manuscript. JK: contributed to the study design, supervision, reviewing and editing of the manuscript. MPIT: contributed to study design, recruitment of trial participants, reviewing and editing of the manuscript. All authors read and approved the final manuscript.

## Data availability {29}

An anonymised version of the data generated and/or analysed during this study are available via the UK Data Archive upon written request, subject to independent review, following the publication of the main findings. A data-sharing agreement would be signed by both parties, outlining the new purpose for which the data are used, conditions of use, appropriate acknowledgement of original funding source, etc.

## Declarations

### Ethics approval and consent to participate {24}

The study has been approved by the London-Stanmore Research Ethics Committee (reference: 21/LO/0896). Informed consent is obtained from all participants in the study.

### Consent for publication {32}

Not applicable—no identifying images or other personal or clinical details of participants are presented here or will be presented in reports of the trial results. The participant information materials and informed consent form are available from the corresponding author on request.

### Competing interests {28}

The authors declare that they have no competing interests and are independent of the contractors.

### Author details

[1]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK. [2]Department of Psychology, University of Bath, Bath, UK. [3]Department of Psychiatry, University of Oxford, Oxford, UK. [4]Department of Informatics, King's College London, London, UK. [5]South London and Maudsley NHS Foundation Trust, London, UK. [6]Kent and Medway Medical School, Canterbury, Kent, UK. [7]Department of Psychological Medicine, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand. [8]College of Community Development and Personal Wellbeing, Otago Polytechnic, Dunedin, New Zealand.

## References

1. Freeman D, Garety PA, Bebbington PE, et al. Psychological investigation of the structure of paranoia in a non-clinical population. Br J Psychiatry. 2005;186(5):427–35. https://doi.org/10.1192/bjp.186.5.427.
2. Bentall RP, Rowse G, Shryane N, et al. The cognitive and affective structure of paranoid delusions. Arch Gen Psychiatry. 2009;66(3):236. https://doi.org/10.1001/archgenpsychiatry.2009.1.
3. Birchwood M, Trower P. The future of cognitive-behavioural therapy for psychosis: not a quasi-neuroleptic. Br J Psychiatry. 2006;188(2):107–8. https://doi.org/10.1192/bjp.bp.105.014985.
4. Garety PA, Kuipers E, Fowler D, Freeman D, Bebbington PE. A cognitive model of the positive symptoms of psychosis. Psychol Med. 2001;31(2):189–95. https://doi.org/10.1017/S0033291701003312.

5.  Savulich G, Freeman D, Shergill S, Yiend J. Interpretation biases in para-
    noia. Behav Ther. 2015;46(1):110–24. https://doi.org/10.1016/j.beth.2014.
    08.002.
6.  Freeman D, Garety PA, Kuipers E, Fowler D, Bebbington PE. A cognitive
    model of persecutory delusions. Br J Clin Psychol. 2002;41(4):331–47.
    https://doi.org/10.1348/014466502760387461.
7.  Wessely S, Buchanan A, Reed A, et al. Acting on delusions. I: prevalence.
    Br J Psychiatry. 1993;163(1):69–76. https://doi.org/10.1192/bjp.163.1.69.
8.  Castle DJ, Phelan M, Wessely S, Murray RM. Which patients with non-
    affective functional psychosis are not admitted at first psychiatric con-
    tact? Br J Psychiatry. 1994;165(1):101–6. https://doi.org/10.1192/bjp.165.1.
    101.
9.  Johnson J. The validity of major depression with psychotic features based
    on a community study. Arch Gen Psychiatry. 1991;48(12):1075. https://
    doi.org/10.1001/archpsyc.1991.01810360039006.
10. Sher L. Manic-depressive illness: bipolar disorders and recurrent depres-
    sion, second edition by Frederick K. Goodwin, M.D., and Kay Redfield
    Jamison, Ph.D. New York, Oxford University Press, 2007, 1,288 pp., $99.00.
    Am J Psychiatry. 2008;165(4):541–2. https://doi.org/10.1176/appi.ajp.2007.
    07121846.
11. Hamner MB, Frueh BC, Ulmer HG, Arana GW. Psychotic features and
    illness severity in combat veterans with chronic posttraumatic stress
    disorder. Biol Psychiatry. 1999;45(7):846–52. https://doi.org/10.1016/
    S0006-3223(98)00301-1.
12. van Os J, Verdoux H, Maurice-Tison S, et al. Self-reported psychosis-like
    symptoms and the continuum of psychosis. Soc Psychiatry Psychiatr
    Epidemiol. 1999;34(9):459–63. https://doi.org/10.1007/s001270050220.
13. Appelbaum PS, Robbins PC, Roth LH. Dimensional approach to
    delusions: comparison across types and diagnoses. Am J Psychiatry.
    1999;156(12):1938–43. https://doi.org/10.1176/ajp.156.12.1938.
14. van der Gaag M, Valmaggia LR, Smit F. The effects of individually tailored
    formulation-based cognitive behavioural therapy in auditory hallucina-
    tions and delusions: A meta-analysis. Schizophr Res. 2014;156(1):30–7.
    https://doi.org/10.1016/j.schres.2014.03.016.
15. Mehl S, Werner D, Lincoln TM. Does cognitive behavior therapy for
    psychosis (CBTp) show a sustainable effect on delusions? A meta-analysis.
    Front Psychol. 2015;6. https://doi.org/10.3389/fpsyg.2015.01450.
16. Burgess-Barr S, Nicholas E, Venus B, et al. International rates of receipt of
    psychological therapy for psychosis and schizophrenia: systematic review
    and meta-analysis. Int J Ment Health Syst. 2023;17(1):8. https://doi.org/10.
    1186/s13033-023-00576-9.
17. Turner DT, van der Gaag M, Karyotaki E, Cuijpers P. Psychological interven-
    tions for psychosis: a meta-analysis of comparative outcome studies. Am
    J Psychiatry. 2014;171(5):523–38. https://doi.org/10.1176/appi.ajp.2013.
    13081159.
18. Elkis H. Treatment-resistant schizophrenia. Psychiatr Clin North Am.
    2007;30(3):511–33. https://doi.org/10.1016/j.psc.2007.04.001.
19. Moritz S, Woodward TS. Metacognitive training in schizophrenia: from
    basic research to knowledge translation and intervention. Curr Opin
    Psychiatry. 2007;20(6):619–25. https://doi.org/10.1097/YCO.0b013e3282
    f0b8ed.
20. Waller H, Freeman D, Jolley S, Dunn G, Garety P. Targeting reason-
    ing biases in delusions: a pilot study of the Maudsley Review Training
    Programme for individuals with persistent, high conviction delusions. J
    Behav Ther Exp Psychiatry. 2011;42(3):414–21. https://doi.org/10.1016/j.
    jbtep.2011.03.001.
21. Hallion LS, Ruscio AM. A meta-analysis of the effect of cognitive bias
    modification on anxiety and depression. Psychol Bull. 2011;137(6):940–58.
    https://doi.org/10.1037/a0024355.
22. Yiend J, Lam CLM, Schmidt N, et al. Cognitive bias modification for para-
    noia (CBM-pa): a randomised controlled feasibility study in patients with
    distressing paranoid beliefs. Psychol Med. 2023;53(10):4614–26. https://
    doi.org/10.1017/S0033291722001520.
23. Yiend J, Trotta A, Meek C, et al. Cognitive Bias Modification for para-
    noia (CBM-pa): study protocol for a randomised controlled trial. Trials.
    2017;18(1):298. https://doi.org/10.1186/s13063-017-2037-x.
24. Hsu, et al. User-Centered Development of STOP (successful treatment
    for paranoia): material development and usability testing for a digital
    therapeutic for paranoia. JMIR Hum Factors. 2023;10:e45453. https://doi.
    org/10.2196/45453.
25. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale
    (PANSS) for schizophrenia. Schizophr Bull. 1987;13(2):261–76. https://doi.
    org/10.1093/schbul/13.2.261.
26. Mathews A, Mackintosh B. Induced emotional interpretation bias and
    anxiety. J Abnorm Psychol. 2000;109(4):602–15.
27. Eysenck MW, Mogg K, May J, Richards A, Mathews A. Bias in interpretation
    of ambiguous sentences related to threat in anxiety. J Abnorm Psychol.
    1991;100(2):144–50. https://doi.org/10.1037/0021-843X.100.2.144.
28. Sturman E. The capacity to consent to treatment and research: a review
    of standardized assessment tools. Clin Psychol Rev. 2005;25(7):954–74.
    https://doi.org/10.1016/j.cpr.2005.04.010.
29. Bush B, Shaw S, Cleary P, Delbanco TL, Aronson MD. Screening for alcohol
    abuse using the cage questionnaire. Am J Med. 1987;82(2):231–5. https://
    doi.org/10.1016/0002-9343(87)90061-1.
30. Imran JB, Richmond RE, Madni TD, et al. Determining suicide risk in
    trauma patients using a universal screening program. J Trauma Acute
    Care Surg. 2018;85(1):182–6. https://doi.org/10.1097/TA.0000000000
    001899.
31. Hettige NC, Bani-Fatemi A, Kennedy JL, De Luca V. Assessing the risk for
    suicide in schizophrenia according to migration, ethnicity and geographi-
    cal ancestry. BMC Psychiatry. 2017;17(1):63. https://doi.org/10.1186/
    s12888-016-1180-3.
32. Mental Capacity Act 2005. Available from: https://www.legislation.gov.uk/
    ukpga/24/9/contents. Last accessed 07 Oct 2024.
33. Fenigstein A, Vanable PA. Paranoia and self-consciousness. J Pers Soc
    Psychol. 1992;62(1):129–38. https://doi.org/10.1037/0022-3514.62.1.129.
34. Aitken RC. Measurement of feelings using visual analogue scales. Proc R
    Soc Med. 1969;62(10):989–93.
35. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasi-
    bility, acceptability, and preliminary efficacy of a smartphone interven-
    tion for schizophrenia. Schizophr Bull. 2014;40(6):1244–53. https://doi.
    org/10.1093/schbul/sbu033.
36. Keetharuth AD, Brazier J, Connell J, et al. Recovering Quality of Life
    (ReQoL): a new generic self-reported outcome measure for use
    with people experiencing mental health difficulties. Br J Psychiatry.
    2018;212(1):42–9. https://doi.org/10.1192/bjp.2017.10.
37. Hartman DE. Wechsler Adult Intelligence Scale IV (WAIS IV): return of the
    gold standard. Appl Neuropsychol. 2009;16(1):85–7. https://doi.org/10.
    1080/09084280802644466.
38. Ray MK, Zachmann AE, Caudill CV, Boggiano MM. Relationship between
    trait suggestibility and eating-related behaviors in overweight and obesity.
    Eat Behav. 2020;37:101380. https://doi.org/10.1016/j.eatbeh.2020.101380.
39. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/
    expectancy questionnaire. J Behav Ther Exp Psychiatry. 2000;31(2):73–86.
    https://doi.org/10.1016/S0005-7916(00)00012-4.
40. Savulich G, Shergill SS, Yiend J. Interpretation biases in clinical paranoia.
    Clin Psychol Sci. 2017;5(6):985–1000. https://doi.org/10.1177/2167702617
    718180.
41. Mallinckrodt CH, Clark WS, David SR. Accounting for dropout bias using
    mixed-effects models. J Biopharm Stat. 2001;11(1–2):9–21. https://doi.
    org/10.1081/BIP-100104194.
42. Freeman D, Loe BS, Kingdon D, et al. The revised Green et al., Paranoid
    Thoughts Scale (R-GPTS): psychometric properties, severity ranges, and
    clinical cut-offs. Psychol Med. 2021;51(2):244–53. https://doi.org/10.1017/
    S0033291719003155.
43. Taher R, Hall C, Bergin A, et al. Developing a process for assessing the
    safety of a digital mental health intervention and gaining regulatory
    approval: a case study and academic's guide. Trials. 2024;25:604. https://
    doi.org/10.1186/s13063-024-08421-1.
44. International Council for Harmonisation of Technical Requirements for
    Pharmaceuticals for Human Use (ICH). Addendum on estimands and
    sensitivity analysis in clinical trials to the guideline on statistical principles
    for clinical trials E9(R1); 2019. Available from: https://database.ich.org/
    sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf. Last accessed
    08 Oct 2024.
45. White IR, Thompson SG. Adjusting for partially missing baseline measure-
    ments in randomized trials. Stat Med. 2005;24(7):993–1007. https://doi.
    org/10.1002/sim.1981.
46. Rubin DB. Multiple Imputation for Nonresponse in Surveys. Wiley. 1987.
    https://doi.org/10.1002/9780470316696.

47.  Carpenter JR, Smuk M. Missing data: a statistical framework for practice. Biom J. 2021;63(5):915–47. https://doi.org/10.1002/bimj.202000196.

48.  Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. J Biopharm Stat. 2013;23(6):1352–71. https://doi.org/10.1080/10543406.2013.834911.

49.  White IR. Strategies for handling missing data in randomised trials. Trials. 2011;12(S1):A59. https://doi.org/10.1186/1745-6215-12-S1-A59.

50.  Downey RG, King CV. Missing data in Likert ratings: a comparison of replacement methods. J Gen Psychol. 1998;125(2):175–91. https://doi.org/10.1080/00221309809595542.

51.  Roth PL, Switzer FS, Switzer DM. Missing data in multiple item scales: a Monte Carlo analysis of missing data techniques. Organ Res Methods. 1999;2(3):211–32. https://doi.org/10.1177/109442819923001.

52.  European Commission. MEDDEV 2.7/3, revision 3: Clinical investigations: serious adverse reporting under directives 90/385/EEC and 93/42/EC. Available from: https://www.medical-device-regulation.eu/wp-content/uploads/2019/05/2_7_3_rev_5_2015_en.pdf. Last accessed 08 Oct 2024.

53.  Health Research Authority. Safety reporting for non-CTIMP studies [updated 09 Aug 2024]. Available from: https://www.hra.nhs.uk/approvals-amendments/managing-your-approval/safety-reporting/. Last accessed 08 Oct 2024.

54.  Parkins, K. Covid-19: Mental health app growth reveals unmet needs. 14 Mar 2022. Available from: https://www.medicaldevice-network.com/features/covid-19-mental-health-app-growth-reveals-unmet-needs/. Last accessed 08 Oct 2024.

55.  Goldberg SB, Lam SU, Simonsson O, Torous J, Sun S. Mobile phone-based interventions for mental health: a systematic meta-review of 14 meta-analyses of randomized controlled trials. PLOS Digit Health. 2022;1(1):e0000002. https://doi.org/10.1371/journal.pdig.0000002.

56.  Lehtimaki S, Martic J, Wahl B, Foster KT, Schwalbe N. Evidence on digital mental health interventions for adolescents and young people: systematic overview. JMIR Ment Health. 2021;8(4):e25847. https://doi.org/10.2196/25847.

57.  Iacobucci G. NICE recommends online therapies to treat depression and anxiety in adults. BMJ. 2023;380:495. https://doi.org/10.1136/bmj.p495.

58.  May CR, Finch T, Ballini L, et al. Evaluating complex interventions and health technologies using normalization process theory: development of a simplified approach and web-enabled toolkit. BMC Health Serv Res. 2011;11(1):245. https://doi.org/10.1186/1472-6963-11-245.

59.  Smith L, Leung WG, Crane B, Parkinson B, Toulopoulou T, Yiend J. Bilingual comparison of Mandarin and English cognitive bias tasks. Behav Res Methods. 2018;50(1):302–12. https://doi.org/10.3758/s13428-017-0871-0.

60.  Rude SS, Valdez CR, Odom S, Ebrahimi A. Negative cognitive biases predict subsequent depression. Cognit Ther Res. 2003;27(4):415–29. https://doi.org/10.1023/A:1025472413805.

61.  Rude SS, Wenzlaff RM, Gibbs B, Vane J, Whitney T. Negative processing biases predict subsequent depressive symptoms. Cogn Emot. 2002;16(3):423–40. https://doi.org/10.1080/02699930143000554.

62.  Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand. 1983;67(6):361–70. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x.

63.  Spurgeon L, James G, Sackley C. The Hospital Anxiety and Depression Scale: a pilot study to examine its latent structure and the link between psychological state and symptom severity in transient ischaemic attack patients. Psychol Health Med. 2016;21(5):632–8. https://doi.org/10.1080/13548506.2015.1074711.

64.  Cameron IM, Crawford JR, Lawton K, Reid IC. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. Br J Gen Pract. 2008;58(546):32–6. https://doi.org/10.3399/bjgp08X263794.

65.  Freeman D, Bird JC, Loe BS, et al. The Dunn Worry Questionnaire and the Paranoia Worries Questionnaire: new assessments of worry. Psychol Med. 2020;50(5):771–80. https://doi.org/10.1017/S0033291719000588.

66.  Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes. 2007;5(1):63. https://doi.org/10.1186/1477-7525-5-63.

67.  Neil ST, Kilbride M, Pitt L, et al. The questionnaire about the process of recovery (QPR): a measurement tool developed in collaboration with service users: corrigendum. Psychosis: psychological, social and integrative approaches. 2010;2(1):88–91. https://doi.org/10.1080/17522430903553495.

68.  Law H, Neil ST, Dunn G, Morrison AP. Psychometric properties of the Questionnaire about the Process of Recovery (QPR). Schizophr Res. 2014;156(2–3):184–9. https://doi.org/10.1016/j.schres.2014.04.011.

69.  Rozental A, Kottorp A, Forsström D, et al. The Negative Effects Questionnaire: psychometric properties of an instrument for assessing negative effects in psychological treatments. Behav Cogn Psychother. 2019;47(5):559–72. https://doi.org/10.1017/S1352465819000018.

## Publisher's Note