

Audio Visual Assisted Therapy Aid for Refractory Auditory Hallucinations (AVATAR) therapy for voice hearers: systematic review and meta-analysis

Felix Oppen, Sebastian Henges, Pawel Weinstein, Dana Arnheim, Laura Fässler, Olivier Percie du Sert, Izabela Stefaniak, Michel Sabé, Louise Birkedal Glenthøj, Neil Thomas, Chih-Sung Liang, Brendon Stubbs and Kerem Böge

Background

Auditory verbal hallucinations (AVHs) are common and distressing symptoms across a range of psychiatric disorders, including schizophrenia spectrum disorders. Audio Visual Assisted Therapy Aid for Refractory Auditory Hallucinations (AVATAR) is an innovative therapeutic approach that facilitates dialogue with a digital avatar representing the voices that patients hear.

Aims

This systematic review and meta-analysis aimed to assess the efficacy, tolerability and acceptability of AVATAR therapy in reducing voice-related symptoms.

Method

Following preregistration, we conducted a systematic review and meta-analysis of controlled trials of AVATAR therapy in samples primarily diagnosed with schizophrenia spectrum disorders. PubMed, CINAHL, Embase, PsycInfo, ClinicalTrials.gov, ISRCTN and Web of Science were searched in March 2025. We assessed bias and certainty with the Cochrane Risk-of-Bias 2 tool and the GRADE approach. Random-effects models were used to synthesise outcomes.

Results

Eight AVATAR trials ($N = 978$) were included. Compared with usual treatment, waitlist and active control groups, AVATAR therapy decreased the primary outcome of AVH severity at

post-treatment (Hedges' $g = -0.40$, 95% CI -0.54 to -0.25) and short-term follow-up (Hedges' $g = -0.25$, 95% CI -0.40 to -0.10). AVH subscales showed small significant effect sizes at post-treatment (frequency: Hedges' $g = -0.38$, 95% CI -0.52 to -0.24 ; distress: Hedges' $g = -0.32$, 95% CI -0.46 to -0.18), which were maintained at short-term follow-up. The certainty of evidence was rated moderate for AVH severity at post-treatment. AVATAR therapy was largely tolerable and acceptable, with adverse events mostly unrelated to the treatment and a comparable drop-out rate to control groups.

Conclusions

Findings suggest that AVATAR therapy is effective at reducing AVH symptoms. Considering heterogeneous control groups and less clear evidence for secondary outcomes and longer follow-ups, further research is warranted.

Keywords

AVATAR; digital intervention; voice hearing; schizophrenia spectrum disorders; virtual reality.

Copyright and usage

© The Author(s), 2026. Published by Cambridge University Press on behalf of Royal College of Psychiatrists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Auditory verbal hallucinations (AVHs), the experience of hearing voices without an external source,¹ can occur across a wide range of psychiatric disorders, including schizophrenia spectrum disorders, affective disorders, borderline personality disorder, post-traumatic stress disorder, as well as in non-clinical populations.² Despite their cross-diagnostic presence, AVHs are most commonly associated with schizophrenia spectrum disorders, where approximately 70% of patients experience them at some point during their lives.³ Phenomenologically, AVHs are highly heterogeneous, varying in content, emotional valence and perceived agency.⁴ They may be personified or non-personified, and experienced as persecutory, abusive, obscene, derogatory, threatening or critical, but also potentially helpful, affirming or inspirational.³

Although not inherently pathological, AVHs can become clinically relevant when they are experienced as intrusive, uncontrollable or malevolent.⁵ In such cases, AVHs are often associated with heightened distress, functional impairment and increased psychopathology.⁶

In the treatment of schizophrenia spectrum disorders, guidelines emphasise a multidisciplinary approach, with antipsychotic medication as a central component.^{7–9} Although antipsychotics are

effective for a substantial proportion of patients,¹⁰ approximately 20–35% do not experience clinically meaningful improvement.^{11–13} Limitations are further underscored by high relapse rates upon medication discontinuation,^{14,15} and the risk of burdensome side-effects.¹⁶ Notably, around 30% of treatment-resistant symptoms involve persistent AVHs,^{17–19} highlighting the urgent need for additional, targeted interventions for individuals who continue to experience significant voice-related distress despite standard pharmacological care.

In addition, psychological interventions such as cognitive-behavioural therapy (CBT) are recommended.^{7–9} CBT has demonstrated small effects on psychotic symptoms in numerous domains, including overall positive²⁰ and AVH-related symptoms.²¹ However, CBT shows relevant limitations including a large number of sessions required for effectiveness.²² More recently, a shift toward symptom-specific approaches has gained momentum, allowing for more personalised treatment strategies.²³ For AVHs, relational therapies have emerged as a promising line of intervention, drawing on the person-like qualities of voices and conceptualising AVHs as embedded within a dynamic, relationship-like framework.²⁴ Among these, the therapy intervention Audio Visual Assisted

Therapy Aid for Refractory Auditory Hallucinations (AVATAR) represents a novel therapeutic development explicitly aimed at improving outcomes for individuals experiencing distressing and persistent voices.²⁵

Initially developed by Julian Leff in 2008,²⁶ AVATAR therapy enables real-time interactions with a digital representation of a person's most dominant AVH. The approach involves the creation of a patient-designed digital avatar displayed either on a screen or through a virtual reality headset. A clinician then animates the avatar by voicing it according to the patient's descriptions, and the avatar's facial and head movements are synchronised to simulate natural speech. This simulated 'face-to-face' interaction with the avatar functions as a form of exposure to anxiety-provoking stimuli,²⁷ with the therapeutic aim of gradually increasing the individual's sense of control, reducing fear-based appraisals and altering the relational dynamics with the voice. By enabling patients to assert themselves and challenge previously threatening voices, AVATAR therapy may reduce the distress associated with AVHs.²⁸ In line with this, maladaptive appraisals of voices, such as beliefs about omnipotence or malevolence, have shown associations with voice-related distress, whereas more positive interpretations were modestly associated with reduced distress.²⁹

Since its creation, AVATAR therapy has been evaluated in several clinical trials.²⁵ Previous reviews have examined AVATAR therapy among virtual reality-based treatments in mental disorders or positive symptoms of schizophrenia spectrum disorders,^{30,31} but they have not provided a comprehensive perspective of AVATAR therapy for AVHs beyond its role within the broader context of virtual reality-based treatments. Two previous systematic reviews and meta-analyses have examined AVATAR trials and have reported promising effects on AVH-related symptoms.^{32,33} However, these were limited by the inclusion of few studies in the meta-analysis, by analysis of short follow-ups and by the lack of secondary aspects such as tolerability and acceptability of the treatment. These aspects are especially relevant for clinical decision-makers and guideline developers in evaluating the real-world utility of emerging therapies. Moreover, further high-quality clinical trials may have since been published, revealing the necessity for an updated review. In light of these limitations, we aimed to comprehensively assess the efficacy of AVATAR-based interventions for AVHs, including both clinical and functional outcomes, as well as their tolerability, acceptability and the overall quality of available evidence.

Method

This systematic review and meta-analysis was preregistered at PROSPERO on 15 March 2025 (identifier CRD420251005545) and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines (see Supplementary Tables 1 and 2 available at <https://doi.org/10.1192/bjo.2026.11014>).³⁴

Search strategy

On 21 March 2025, the publication databases Embase, PubMed and PsycINFO, as well as the grey literature and trial registration databases Web of Science Core Collection, CENTRAL, ClinicalTrials.gov and ISRCTN, were searched for relevant studies, to minimise the potential of publication bias. Indexing and general terms were applied in the search (see Supplementary Table 3).³⁵ References of prior reviews and included articles were checked for additional studies, and experts were contacted for knowledge of further suitable trials. Duplicate removal and record screening were performed in the systematic review program Rayyan.³⁶ F.O. and

S.H. independently screened titles and abstracts for full-text screening eligibility with hierarchical criteria (see Supplementary Table 4),³⁷ before independently completing the full-text screening in a spreadsheet. Corresponding authors of trial registrations and missing data were contacted weekly over the period of 4 weeks to request missing information. For disagreements during screenings, a third reviewer (K.B.) was consulted for discussion. Interrater reliability in the metric of Cohen's κ was calculated using the DeltaMAN package for both screening stages.^{38,39}

Inclusion criteria

To be included, studies had to be published or unpublished, randomised controlled trials or controlled trials, investigating the efficacy of AVATAR or equivalent therapy in treating samples of participants aged ≥ 16 years who reported AVHs and were diagnosed with a schizophrenia spectrum disorder in $>50\%$ of cases, compared to any active or passive control group. Finally, studies had to report one of the following outcomes of interest.

The primary outcome of this study was the between-group difference of severity of voice symptoms at post-treatment measured with the Psychotic Symptom Rating Scales for Auditory Hallucinations (PSYRATS-AH⁴⁰). Secondary outcomes included further voice-related (frequency and distress in the PSYRATS-AH), clinical (positive, negative, total psychotic, depressive and anxiety symptoms) and functional outcomes (social functioning and quality of life), as well as their follow-up assessments. We also indexed acceptability by examining within- and between-group drop-out rates, and tolerability from reports of treatment-related adverse events and symptom exacerbations. For multiple non-AVH measurements per study and outcome, the more primary measurement was included into the quantitative analysis (per study definition or earlier reference as an outcome). Follow-up measurements were categorised as short (12–23 weeks), medium (24–51 weeks) and long term (≥ 52 weeks) as a pragmatic approach reflecting common follow-up durations selected in meta-analyses of psychological intervention trials in schizophrenia spectrum disorders.^{41,42}

We examined both treatment and study drop-out. Treatment drop-out was identified by any participant who did not complete the treatment post-randomisation.⁴³ The criteria for having finished a treatment followed that of each study. In contrast, study drop-out was characterised by participants lost to post-treatment assessments regardless of reason.

Data extraction, synthesis and effect sizes

Data was independently extracted by F.O. and S.H., and cross-checked by D.A. The data extraction started on 2 April 2025. Meta-analysis was performed when at least two effect sizes could be pooled per analysis. Although AVATAR therapy approaches are similar, we anticipated differences in control groups, leading us to the use of random-effects models, which allow for heterogeneous treatment effects.⁴⁴

The metric of Hedges' g was used to report the standardised mean difference for continuous outcomes and aggregate outcomes of overlapping constructs using different scales.⁴⁵ Negative effect sizes described outcomes with a lower mean score in the AVATAR groups relative to the control groups, with positive values indicating the opposite. Values of 0.2, 0.5 and 0.8 were considered the thresholds for low, medium and large effect sizes, respectively.⁴⁶ Proportion and risk ratio were calculated for within- and between-group drop-out effect sizes. Drop-out proportions were calculated with generalised linear-mixed models.⁴⁷ Risk ratios >1 portrayed an increased risk in AVATAR relative to control groups. To adjust for zero-case studies, 0.5 cases of drop-out were added to each

study's risk table containing at least one instance of zero drop-outs in either group.⁴⁸ For three-arm studies, the sample sizes of the control groups were halved to include both AVATAR groups, per the suggestions in the Cochrane Handbook.⁴⁹

The presence of heterogeneity was tested with Cochran's *Q*-test and interpreted as Higgins' *I*²-statistic.^{50,51} Percentages of 25%, 50% and 75% were considered low, moderate and high levels of heterogeneity, respectively. All analyses were performed in R version 2025.05.1+513 in RStudio for macOS, using the metafor and Tidyverse packages, and were visualised with forest plots.^{52–54} To exploratively assess the robustness of effects on AVH symptoms, subgroup analyses of studies with low and high risks of bias were performed. Additionally, for AVH symptoms, analyses by control group (treatment as usual versus treatment as usual plus waitlist versus active control group) were performed with subgroup analyses and mixed-effects meta-regression. Results with significant heterogeneity were examined with jackknife analyses.⁵⁵ An α -level of 0.05 was considered for all analyses.

Bias assessments, publication bias and quality of evidence

F.O. and S.H. independently assessed risk of bias with the Revised Cochrane Risk-of-Bias Tool for Randomized Trials (RoB-2⁵⁶), and differences were discussed with a third reviewer, K.B. Bias was rated under consideration of the primary outcome of AVH severity. The risk of bias was visualised using the robvis app.⁵⁷ Publication bias may occur when articles remain unpublished because of unwanted study results.⁵⁸ For main analyses, funnel plots were plotted for its assessment and asymmetry was tested with Egger's tests.⁵⁹ Asymmetric analyses were eligible to be corrected with the trim-and-fill procedure.⁶⁰ F.O. characterised the certainty of meta-analytic evidence according to the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) criteria, using GRADEpro online software.^{61,62}

Results

Study selection

The literature search identified 556 records (see the PRISMA flow diagram in Supplementary Fig. 1). After duplicate removal and dual screenings, a total of seven peer-reviewed articles and one letter to the editor were included. We contacted authors of trial registrations and conference abstracts, although no ongoing trials were able to provide includable information concerning relevant outcomes for this study. Abstract and full-text screening exclusion reasons can be viewed in Supplementary Tables 4 and 5. The original AVATAR trial included a small number of participants under 16 years of age,²⁵ formally conflicting with our inclusion criteria. Given this small proportion and comparable mean age and standard deviation to other included samples, the study was retained. The title and abstract as well as full-text screening interrater reliabilities revealed Cohen's $\kappa = 0.95$ and 0.93 , respectively, which can be considered satisfactory.⁶³

Study and sample characteristics

The eight randomised and controlled studies included nine relevant comparisons (Garety et al.⁶⁴ included two AVATAR groups) and $n = 554$ participants in the AVATAR therapy groups and $n = 424$ in the control groups. Of note is that three studies were partial crossover trials,^{25,65,66} with only the period before crossover included in the analyses. The post-treatment measurement varied from 7 to 16 weeks (mean 9.88 weeks). Each study that performed a

follow-up assessment did so at 12–13-week follow-up, and one study additionally assessed at 24- and 52-week follow-ups.⁶⁷ All participants received treatment as usual, typically including antipsychotic medication. Control groups consisted of treatment as usual alone,⁶⁴ treatment as usual plus waitlist,^{25,65,66} or treatment as usual with an additional active control condition (CBT,^{67,68} supportive counselling,²⁶ enhanced treatment as usual with supportive counselling sessions⁶⁹). The studies with an additional active control condition were categorised as active control groups in the subgroup analysis. The range of treatment durations was 6 to 9 weeks or sessions except for AVATAR-EXT,⁶⁴ an extension of AVATAR in 12 weekly sessions. Samples consisted of persons with schizophrenia spectrum disorders reporting treatment-resistant, chronic or persisting AVHs, although three studies included small subsamples of other mental disorders.^{26,64,66} Demographic and clinical outcomes, as well as specific study characteristics, can be found in Table 1. Additional information (protocols, criteria, settings, diagnosis distributions, participant ages and funding) is provided in Supplementary Table 6, whereas Supplementary Table 7 describes additional information received from study authors.

AVATAR therapy protocols

The treatment protocols showed differences. Treatment began with the creation of a digital avatar, except in Stefaniak et al.⁶⁶ which used a standardised avatar. Although some protocols enabled viewing the avatar via a virtual reality headset,^{65,67–69} others presented the avatar on a computer screen,^{25,26,64,66} and one used full-face and body motion capture of the therapist to increase the feeling of presence and immersion.⁶⁸ The participants usually interacted with an avatar voiced by the therapist in real time, although in Stefaniak et al.⁶⁶ a pre-programmed avatar was used, which the therapist and participant both interacted with.

Post-creation sessions included a therapeutic session preparation phase, followed by exposure to distressing utterances of the avatar, during which the participant was encouraged to respond assertively.²⁶ As the treatment progressed, the avatar's verbalisations gradually became less abusive and more supportive to reflect the participants' increase in control. Later sessions were designated for future outlook and relapse prevention. As an extension of this framework, AVATAR-EXT provided six additional sessions, where participants discussed the trauma, marginalisation and social exclusion background of their AVHs with the therapist before avatar exposure.⁶⁴ Half of the studies provided recordings of the sessions, which could be listened to as homework to increase transferral to daily life.^{25,26,64,68}

Risk of bias

The risk-of-bias assessments are displayed in Fig. 1, and the results of the bias assessment questions can be found in Supplementary Table 8. Overall, two studies were rated with low risk of bias,^{64,69} and the remaining six studies high overall risk of bias.^{25,26,65–68} Importantly, the two studies rated as having low risk of bias accounted for the majority (63%) of participants.

Meta-analysis and systematic review results

Although control groups varied across studies, a main analysis of all studies was performed. Treatments for schizophrenia spectrum disorders differ depending on the setting, availability and severity of symptoms.^{70–72} Consequently, this approach provided greater external validity, but may have induced heterogeneity. Subgroup analyses by specific control group type were conducted to

Table 1 Study characteristics

Study and country	Experimental and control group conditions	N _{Experimental group, control group}	Duration	Sample type and diagnoses	Outcome measures of interest	Assessments (after T0)	Study design
Craig et al, 2018 (UK) ²⁶	AVATAR + TAU Supportive counselling + TAU	75 v. 75	Six weekly sessions at 50 min, extendable up to ten sessions	Treatment-resistant AVHs; schizophrenia, schizoaffective disorder, bipolar disorder, unspecified psychosis, depression with psychotic symptoms	Voice severity, frequency, distress, positive symptoms, negative symptoms, depression, quality of life	T0 T1 = 12 weeks T2 = 24 weeks	RCT
Dellazizzo et al, 2021 (Canada) ⁶⁷	AVATAR (virtual reality) + TAU CBT + TAU	37 v. 37	Nine weekly sessions at 60 min	Treatment-resistant AVHs; schizophrenia, schizoaffective disorder	Voice severity, frequency, distress, total psychotic symptoms, positive symptoms, negative symptoms, depression, quality of life	T0 T1 = 9 weeks T2 = 22 weeks T3 = 35 weeks T4 = 61 weeks	RCT
Garety et al, 2024 (UK) ⁶⁴	AVATAR extended + TAU AVATAR brief + TAU TAU	114 116 115	12 sessions at 45–60 min Six sessions at 45–60 min	AVHs persisting for more than 6 months; schizophrenia, persistent delusional disorders, acute and transient psychotic disorders, induced delusional disorder, schizoaffective disorders, other non-organic psychotic disorders, unspecified non-organic psychosis, bipolar affective disorder, severe depressive episode with psychotic symptoms	Voice distress, frequency, total severity, anxiety, depression	T0 T1 = 16 weeks T2 = 28 weeks	RCT
Leff et al, 2013 (UK) ²⁵	AVATAR + TAU TAU + waitlist	14 v. 12	7 weeks, six sessions at 30 min	Treatment-resistant AVHs; schizophrenia	Voice severity, depression, voice frequency, distress	T0 T1 = 7 weeks	RPCT
Liang et al, 2022 (China) ⁶⁸	AVATAR (virtual reality + motion capture) + TAU CBT + TAU	32 v. 33	Seven to nine weekly sessions at 60 min	Treatment-resistant AVHs; schizophrenia	Voice severity, distress, frequency, total psychotic symptoms, positive symptoms, negative symptoms, quality of life, depression, anxiety	T0 T1 = 7–9 weeks T2 = 19–21 weeks	RCT
Percie du Sert et al, 2018 (Canada) ⁶⁵	AVATAR (virtual reality) + TAU TAU + waitlist	12 v. 7	Seven weekly sessions at 45 min, extendable up to 11 sessions	Treatment-resistant AVHs; schizophrenia, schizoaffective disorder	Voice severity, frequency, distress, total psychotic symptoms, positive symptoms, negative symptoms, depression, quality of life	T0 T1 = 7 weeks	RPCT
Smith et al, 2025 (Denmark) ⁶⁹	AVATAR (virtual reality) + TAU Enhanced TAU (supportive counselling)	140 v. 131	12 weeks, seven weekly to bi-weekly sessions at 60 min; two booster sessions after post-treatment	Treatment-resistant AVHs; schizophrenia, persistent delusional disorders, schizoaffective disorder, other non-organic psychotic disorders	Voice severity, frequency, distress, social functioning, positive symptoms, negative symptoms, depression	T0 T1 = 12 weeks T2 = 24 weeks	RCT
Stefaniak et al, 2019 (Poland) ⁶⁶	AVATAR + TAU TAU + waitlist	14 v. 14	Eight sessions at 50 min	Chronic AVHs; schizophrenia; paranoid personality disorder	Positive symptoms, voice severity, frequency, distress	T0 T1 = not reported (likely 8 weeks)	RPCT
N, sample size; T0, baseline; AVATAR, Audio Visual Assisted Therapy Aid for Refractory Auditory Hallucinations; TAU, treatment as usual; AVH, auditory verbal hallucinations; T1, post-treatment; T2, follow-up; RCT, randomised controlled trial; RPCT, randomised partial crossover trial. Sample sizes given at baseline.							

		Risk-of-bias domains					
		D1	D2	D3	D4	D5	Overall
Study	Craig et al, 2018 ²⁶	+	+	×	+	+	×
	Dellazizzo et al, 2021 ⁶⁷	+	+	×	×	-	×
	Garety et al, 2024 Ext. / Brf. ⁶⁴	+	+	+	+	+	+
	Leff et al, 2013 ²⁵	+	+	×	+	-	×
	Liang et al, 2022 ⁶⁸	-	+	×	+	-	×
	Percie du Sert et al, 2018 ⁶⁵	×	+	×	×	-	×
	Smith et al, 2025 ⁶⁹	+	+	+	+	+	+
Stefaniak et al, 2019 ⁶⁶	×	+	×	×	-	×	
		Domains:					Judgement
		D1: Bias arising from the randomisation process.					×
		D2: Bias owing to deviations from intended intervention.					-
		D3: Bias owing to missing outcome data.					+
		D4: Bias in measurement of the outcome.					
		D5: Bias in selection of the reported result.					

Fig. 1 Risk-of-bias assessments. Ext., extended; Brf., brief.

Table 2 Meta-analytic results for auditory verbal hallucination-related outcomes

Outcome	Timeframe/analysis	Comparisons	N _{Experimental group, control group}	I ² %	I ² p-value	Hedges' g [95% CI]	Effect p-value	GRADE rating ^a
AVH severity	T1 (post-treatment)	9	462 v. 352	0	0.23	-0.40 [-0.54, -0.25]	<0.001***	Moderate
	T1 Publication bias correction	11	Not applicable	0	0.04*	-0.35 [-0.49, -0.22]	<0.001***	Not applicable
	T2 (short-term follow-up)	6	418 v. 319	0	0.99	-0.25 [-0.40, -0.10]	<0.001***	Low
	T3 (medium-term follow-up)	1	13 v. 20	Not applicable	Not applicable	-0.30 [-1.00, 0.40]	0.40	Very low
	T4 (long-term follow-up)	1	13 v. 17	Not applicable	Not applicable	-0.26 [-0.99, 0.46]	0.48	Very low
AVH frequency	T1 (post-treatment)	9	462 v. 352	0	0.13	-0.38 [-0.52, -0.24]	<0.001***	High
	T2 (short-term follow-up)	6	421 v. 320	0	0.61	-0.34 [-0.49, -0.19]	<0.001***	Low
	T3 (medium-term follow-up)	1	13 v. 20	Not applicable	Not applicable	-0.33 [-1.03, 0.37]	0.36	Very low
	T4 (long-term follow-up)	1	13 v. 17	Not applicable	Not applicable	-0.40 [-1.12, 0.33]	0.29	Very low
AVH distress	T1 (post-treatment)	9	462 v. 352	0	0.09	-0.32 [-0.46, -0.18]	<0.001***	Low
	T1 Publication bias correction	11	Not applicable	0	0.005**	-0.28 [-0.41, -0.14]	<0.001***	Not applicable
	T2 (short-term follow-up)	6	419 v. 320	0	0.94	-0.20 [-0.35, -0.06]	0.007**	Very low
	T3 (medium-term follow-up)	1	13 v. 20	Not applicable	Not applicable	-0.21 [-0.91, 0.49]	0.55	Very low
	T4 (long-term follow-up)	1	13 v. 17	Not applicable	Not applicable	-0.18 [-0.90, 0.54]	0.62	Very low

N, sample size; I², heterogeneity statistic; GRADE, Grading of Recommendations, Assessment, Development and Evaluation; AVH, auditory verbal hallucinations; T1, post-treatment; T2, follow-up at 12 weeks; T3, follow-up at 24 weeks; T4, follow-up at 52 weeks.
 Negative effect sizes described lower scores in intervention groups.
 a. GRADE ratings⁷³: High: a lot of confidence that the true effect lies close to that of the estimated effect. Moderate: moderate confidence in the estimated effect. The true effect is likely to be close to the estimated effect, but there is a possibility that it is substantially different. Low: limited confidence in the estimated effect. The true effect might be substantially different from the estimated effect. Very low: very little confidence in the estimated effect. The true effect is likely to be substantially different from the estimated effect.
 *p < 0.05; **p < 0.01; ***p < 0.001.

complement this approach. Forest plots of all secondary outcomes can be found in Supplementary Figs 2–6, the assessments of the certainty of evidence according to GRADE criteria can be found in Supplementary Table 9 and information on which scales were included can be found in Supplementary Table 10.

Primary and secondary AVH outcomes: AVH severity, frequency and distress

Table 2 presents the meta-analytic results for voice-related outcomes alongside the corresponding certainty of evidence, and Table 3 shows subgroup and moderation analyses by subgroup and bias. Each study assessed AVH severity, frequency and distress at pre- and post-treatment. Short-term 12- to 13-week follow-up assessments were performed in all but three studies.^{25,65,66} The pooled effect size on AVH severity at post-treatment was small, homogeneous and favoured AVATAR (Hedges' $g = -0.40$, 95% CI -0.54 to -0.25 , $p < 0.001$). This effect was moderated by control group ($p = 0.01$), although each subgroup favoured AVATAR.

At short-term follow-up, the effect size remained small (Hedges' $g = -0.25$, 95% CI -0.40 to -0.10 , $p < 0.001$) and homogeneous. This effect was not moderated by control group ($p = 0.82$), but only the active control subgroup significantly favoured AVATAR. Medium- and long-term effect sizes were derived from only one study, and were small and non-significant. Studies rated with a low risk of bias revealed similar results at both post-treatment and short-term follow-up. Forest plots for AVH severity are displayed in Fig. 2. The certainty of evidence was rated moderate for post-treatment and low for the short-term-follow-up, whereas medium- and long-term follow-ups were rated very low.

For AVH frequency, the pooled effect size at post-treatment was small, favoured AVATAR (Hedges' $g = -0.38$, 95% CI -0.52 to -0.24 , $p < 0.001$) and was homogeneous. Of the control group subgroup analyses, all but the AVATAR versus the treatment as usual plus waitlist group favoured AVATAR. The effect size at short-term follow-up remained small (Hedges' $g = -0.34$, 95% CI -0.49 to -0.19 , $p < 0.001$), significantly favoured AVATAR across subgroups and was homogeneous. Medium- and long-term

Table 3 Meta-analytic subgroup and moderator results for auditory verbal hallucination-related outcomes

Outcome	Timeframe/analysis	Comparisons	N _{Experimental group, control group}	I ² %	I ² p-value	Hedges' g [95% CI]	Effect p-value
AVH severity	T1 Low bias	3	311 v. 205	0	0.87	-0.32 [-0.50, -0.14]	<0.001***
	T1 High bias	6	151 v. 147	36	0.14	-0.59 [-0.91, -0.27]	<0.001***
	T1 TAU control	2	196 v. 103	0	0.86	-0.36 [-0.60, -0.12]	0.003**
	T1 TAU + waitlist control	3	29 v. 25	0	0.75	-1.25 [-1.84, -0.67]	<0.001***
	T1 Active control	4	237 v. 224	0	0.75	-0.33 [-0.52, -0.15]	<0.001***
	T1 Control group comparison	9	462 v. 352	0	0.94	Not applicable	0.01*
	T2 Low bias	3	304 v. 203	0	0.99	-0.22 [-0.40, -0.04]	0.02*
	T2 High bias	3	114 v. 116	0	0.93	-0.31 [-0.57, -0.05]	0.02*
	T2 TAU control	2	194 v. 104	0	0.91	-0.23 [-0.47, 0.01]	0.06
	T2 Active control	4	224 v. 215	0	0.93	-0.26 [-0.45, -0.08]	0.006**
	T2 Control group comparison	6	418 v. 319	0	0.98	Not applicable	0.82
	T2 Control group comparison	6	418 v. 319	0	0.98	Not applicable	0.82
AVH frequency	T1 Low bias	3	311 v. 205	0	0.74	-0.34 [-0.52, -0.16]	<0.001***
	T1 High bias	6	151 v. 147	58	0.04*	-0.46 [-0.86, -0.06]	0.02*
	T1 TAU control	2	196 v. 103	0	0.59	-0.39 [-0.63, -0.15]	0.002**
	T1 TAU + waitlist control	3	29 v. 25	62	0.07	-0.79 [-1.73, 0.15]	0.10
	T1 Active control	4	237 v. 224	30	0.21	-0.32 [-0.55, -0.09]	0.007**
	T1 Control group comparison	9	462 v. 352	13	0.12	Not applicable	0.32
	T2 Low bias	3	307 v. 204	0	0.54	-0.34 [-0.52, -0.16]	<0.001***
	T2 High bias	3	114 v. 116	0	0.31	-0.33 [-0.59, -0.07]	0.01*
	T2 TAU control	2	194 v. 104	0	0.46	-0.41 [-0.65, -0.17]	<0.001***
	T2 Active control	4	227 v. 216	0	0.48	-0.29 [-0.48, -0.11]	0.002**
	T2 Control group comparison	6	421 v. 320	0	0.55	Not applicable	0.45
	T2 Control group comparison	6	421 v. 320	0	0.55	Not applicable	0.45
AVH distress	T1 Low bias	3	311 v. 205	0	0.85	-0.26 [-0.44, -0.08]	0.004**
	T1 High bias	6	151 v. 147	65	0.03*	-0.55 [-0.99, -0.10]	0.02*
	T1 TAU control	2	196 v. 103	0	0.86	-0.31 [-0.55, -0.07]	0.01*
	T1 TAU + waitlist control	3	29 v. 25	33	0.23	-1.13 [-1.85, -0.42]	0.002**
	T1 Active control	4	237 v. 224	0	0.42	-0.25 [-0.44, -0.07]	0.007**
	T1 Control group comparison	9	462 v. 352	0	0.44	Not applicable	0.02*
	T2 Low bias	3	305 v. 204	0	0.94	-0.17 [-0.35, 0.01]	0.06
	T2 High bias	3	114 v. 116	0	0.67	-0.27 [-0.53, -0.01]	0.04*
	T2 TAU control	2	194 v. 104	0	0.99	-0.14 [-0.38, 0.10]	0.24
	T2 Active control	4	225 v. 216	0	0.83	-0.24 [-0.43, -0.05]	0.01*
	T2 Control group comparison	6	419 v. 320	0	0.93	Not applicable	0.53
	T2 Control group comparison	6	419 v. 320	0	0.93	Not applicable	0.53

N, sample size; I², heterogeneity statistic; AVH, auditory verbal hallucinations; T1, post-treatment; TAU, treatment as usual; T2, follow-up at 12–23 weeks; T3, follow-up at 24–51 weeks; T4, follow-up at 52 weeks.
 Negative effect sizes described lower scores in intervention groups.
 *p < 0.05; **p < 0.01; ***p < 0.001.

follow-up effect sizes were derived from one study and were small and non-significant. Studies rated with a low risk of bias revealed similar results at both time points. The certainty of evidence was rated high for post-treatment and low for short term follow-up, whereas medium- and long-term follow-ups were rated very low.

For AVH distress, the pooled effect size at post-treatment was medium, significant (Hedges' $g = -0.32$, 95% CI -0.46 to -0.18 , $p < 0.001$) and homogeneous. This effect was moderated by control group ($p = 0.02$), but each of the subgroups significantly favoured AVATAR. The pooled small effect size at short-term follow-up favoured AVATAR (Hedges' $g = -0.20$, 95% CI -0.35 to -0.06 , $p = 0.007$) and was homogeneous. Only the active control subgroup favoured AVATAR at short-term follow-up. Medium- and long-term follow-up effect sizes were derived from one study and were negligible to small and non-significant. Studies rated with a low risk of bias revealed similar results at post-treatment, but were non-significant at short-term follow-up. The certainty of evidence was rated low for post-treatment and very low for short-term follow-up, whereas medium- and long-term follow-ups were rated very low.

Secondary outcomes: clinical and functional outcomes

Table 4 presents the results of meta-analyses performed for secondary clinical outcomes. PSYRATS subscales measured by one study⁶⁴ were aggregated according to the recommendations of Borenstein,⁷⁴ assuming a correlation of 0.34 between delusions and hallucinations.⁷⁵ For positive symptoms, small and negligible effect sizes significantly favoured AVATAR at post-treatment and

short-term follow-up, respectively. In contrast, effect sizes for total psychotic and negative symptoms did not significantly favour either group at either time point. Small effect sizes significantly favoured AVATAR for anxiety and depressive symptoms at post-treatment, but remained significant only for anxiety symptoms at short-term follow-up. No significant between-group effects were observed for quality of life at either time point. Social functioning was measured by only one study,⁶⁹ and was therefore not meta-analytically aggregated. The removal of a singular outlier⁶⁷ at post-treatment, specifically for quality of life, reduced heterogeneity without altering the conclusion of the statistical tests. Medium- and long-term follow-up assessments were derived from one study and were non-significant for each outcome.

Secondary outcomes: tolerability and acceptability

Each study reported tolerability aspects concerning treatment-related adverse events or exacerbations, except for Leff et al.²⁵ Adverse events and exacerbations were non-standardised, making narrative review necessary. In Percie du Sert et al,⁶⁵ 1 participant out of 12 received additional counselling during early treatment, because of transient symptom exacerbations. Additionally, each of the four participants who dropped out did so because of anxiety symptom exacerbations in early virtual reality sessions. Three studies stated that any adverse events were not attributable to AVATAR or the control group.^{26,67,68} In Garety et al,⁶⁴ 1 case of hospital admission in AVATAR (out of 116) and 4 cases in AVATAR-EXT (out of 114) could not be ruled out from possibly being attributable to the treatments. In Smith et al,⁶⁹ of 140

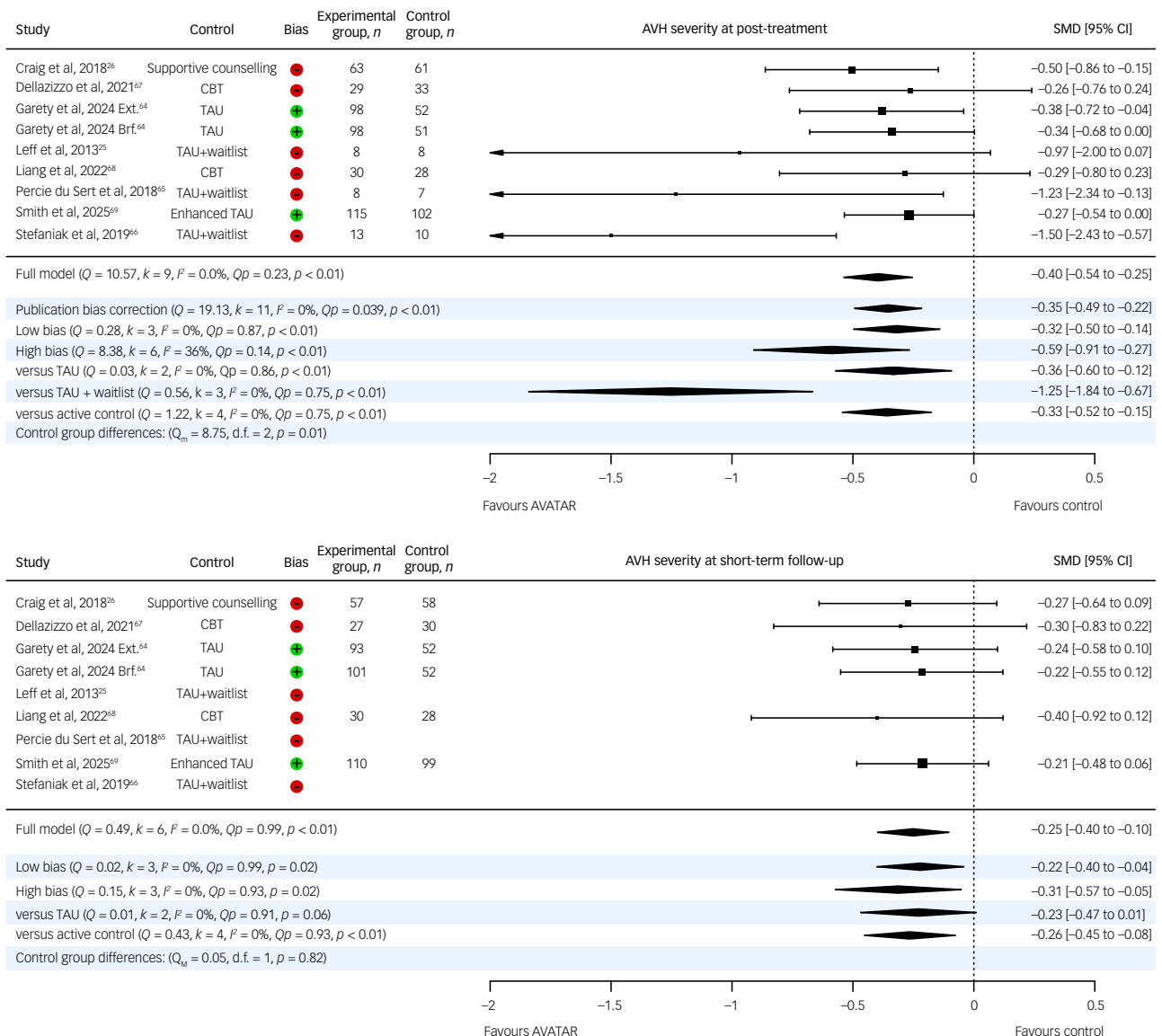


Fig. 2 Forest plot showing meta-analytic results for AVH severity at post-treatment and follow-up. Negative SMDs portray smaller means in AVATAR compared with control groups. ⊕, Low bias; ⊖, high bias; AVATAR, Audio Visual Assisted Therapy Aid for Refractory Auditory Hallucinations; AVH, auditory verbal hallucination; Brf., brief; CBT, cognitive-behavioural therapy; Ext., extended; I^2 , Higgins' heterogeneity statistic; k , number of comparisons; n , sample size; Q , Cochran's Q -statistic; Qp , Cochran's Q p -value; SMD, standardised mean difference in Hedges' g ; TAU, treatment as usual.⁶⁶

participants, five cases of hospital admission and one case of self-harm were considered potentially related to the AVATAR treatment because of worsening of AVHs. AVH symptom increases occurred in 52 participants during early exposure, which gradually declined over the course of the therapy. Furthermore, 40 participants reported needing additional time to manage anxiety during virtual reality immersion. Finally, Stefaniak et al⁶⁶ reported one case of hospital admission among 14 in the AVATAR group, although causal connections to the treatment were not reported.

Each study reported on treatment drop-outs in AVATAR and five reported rates in control groups. As can be seen in Table 5, the overall aggregated proportion of treatment drop-out in AVATAR was 24% with moderate heterogeneity. In control groups, the aggregated proportion of treatment drop-out was 18%, with moderate heterogeneity. No single outlier was responsible for the observed heterogeneity. In terms of study drop-out, both intervention and control groups showed aggregated proportions of 16%. Comparative analyses showed no significant difference in

risk between AVATAR and control groups for both treatment and study drop-out (risk ratio of 1.01 in both cases).

Publication bias

The Egger's test for AVH severity at post-treatment was significant ($p = 0.009$), indicating an asymmetric funnel plot.⁵⁹ Examination of the corresponding funnel plot revealed the notable impact of the large effect size in Stefaniak et al.⁶⁶ Trim-and-fill procedures revealed a corrected small effect size favouring AVATAR (Hedges' $g_{corrected} = -0.35$, 95% CI -0.49 to -0.22 , $p < 0.001$), with two studies estimated to be missing. The Egger's test of AVH distress ($p = 0.02$) at post-treatment was significant. Examination of the corresponding funnel plot revealed the impact of the large effect sizes in Stefaniak et al⁶⁶ and Percie du Sert et al.⁶⁵ Trim-and-fill procedures revealed a corrected small effect size favouring AVATAR (Hedges' $g_{corrected} = -0.28$, 95% CI -0.41 to -0.14 , $p < 0.001$), with two studies estimated to be missing. Funnel plots

Table 4 Meta-analytic results for clinical and functional outcomes

Outcome	Timeframe/analysis	Comparisons	$N_{\text{Experimental group, control group}}$	I^2 %	I^2 p-value	Hedges' g [95% CI]	Effect p-value
Total psychotic symptoms	T1 (post-treatment)	3	67 v. 68	0	0.61	-0.20 [-0.54, 0.14]	0.26
	T2 (short-term follow-up)	2	57 v. 58	0	0.70	-0.17 [-0.54, 0.19]	0.35
	T3 (medium-term follow-up)	1	13 v. 20	Not applicable	Not applicable	0.26 [-0.45, 0.96]	0.47
	T4 (long-term follow-up)	1	13 v. 17	Not applicable	Not applicable	0.00 [-0.72, 0.72]	1
Positive symptoms	T1 (post-treatment)	8	442 v. 341	0	0.47	-0.29 [-0.43, -0.16]	<0.001***
	T2 (short-term follow-up)	6	407.5 v. 315.5 ^a	0	0.64	-0.16 [-0.30, -0.03]	0.02*
	T3 (medium-term follow-up)	1	13 v. 20	Not applicable	Not applicable	0.29 [-0.41, 0.99]	0.42
	T4 (long-term follow-up)	1	13 v. 17	Not applicable	Not applicable	0.22 [-0.50, 0.95]	0.55
Negative symptoms	T1 (post-treatment)	5	245 v. 231	0	0.96	0.09 [-0.09, 0.27]	0.31
	T2 (short-term follow-up)	4	227 v. 216	0	0.90	0.00 [-0.19, 0.19]	0.99
	T3 (medium-term follow-up)	1	13 v. 20	Not applicable	Not applicable	0.14 [-0.56, 0.84]	0.69
	T4 (long-term follow-up)	1	13 v. 17	Not applicable	Not applicable	0.07 [-0.65, 0.79]	0.85
Depressive symptoms	T1 (post-treatment)	8	433 v. 332	46	0.06	-0.25 [-0.47, -0.04]	0.02*
	T2 (short-term follow-up)	6	400 v. 310	21	0.30	-0.13 [-0.30, 0.05]	0.15
	T3 (medium-term follow-up)	1	13 v. 20	Not applicable	Not applicable	0.70 [-0.02, 1.42]	0.06
	T4 (long-term follow-up)	1	13 v. 17	Not applicable	Not applicable	0.31 [-0.42, 1.03]	0.41
Anxiety symptoms	T1 (post-treatment)	4	272 v. 182	0	0.60	-0.45 [-0.65, -0.26]	<0.001***
	T2 (short-term follow-up)	4	260 v. 180	0	0.56	-0.24 [-0.43, -0.05]	0.01*
Quality of life	T1 (post-treatment)	4	128 v. 127	82	0.007**	0.33 [-0.33, 0.99]	0.33
	T1 Removal of outlier ⁶⁷	3	99 v. 94	58	0.09	0.52 [-0.00, 1.05]	0.05
	T2 (short-term follow-up)	3	114 v. 113	0	0.41	0.15 [-0.11, 0.42]	0.25
	T3 (medium-term follow-up)	1	13 v. 20	Not applicable	Not applicable	-0.31 [-1.01, 0.39]	0.39
Social functioning	T4 (long-term follow-up)	1	13 v. 17	Not applicable	Not applicable	-0.40 [-1.13, 0.33]	0.28
	T1 (post-treatment)	1	114 v. 102	Not applicable	Not applicable	-0.04 [-0.31, 0.22]	0.74
	T2 (short-term follow-up)	1	112 v. 99	Not applicable	Not applicable	0.04 [-0.23, 0.31]	0.77

I^2 , heterogeneity statistic; T1, post-treatment; T2, follow-up at 12 weeks; T3, follow-up at 24 weeks; T4, follow-up at 52 weeks; AVATAR, Audio Visual Assisted Therapy Aid for Refractory Auditory Hallucinations.
 Negative effect sizes described lower means in AVATAR groups.
 a. Decimals are because of aggregation of subscales.
 * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 5 Meta-analytic drop-out results

Outcome	Timeframe/analysis	k	$N_{\text{Experimental group, control group}}$	I^2 %	I^2 p-value	Effect size [95% CI]	Effect size p-value
Attrition (treatment drop-out)	Risk ratio	5	170 v. 166	49	0.10	1.01 [0.43–2.35]	0.98
	Proportion in experimental group	9	554	73	<0.001***	Proportion: 24% [0.17–0.33]	<0.001***
	Proportion in control group	5	166	59	0.005**	Proportion: 18% [0.09–0.32]	<0.001***
Attrition (study drop-out)	Risk ratio	9	554 v. 424	0	0.45	Risk ratio: 1.01 [0.75–1.35]	0.96
	Proportion in experimental group	9	554	0	0.10	Proportion: 16% [0.14–0.20]	<0.001***
	Proportion in control group	9	424	15	0.13	16% [0.12–0.21]	<0.001***

N , sample size; I^2 , heterogeneity statistic. Negative effect sizes described lower scores in intervention groups whereas risk ratios above 1 describe an increased risk in intervention groups.
 ** $p \leq 0.01$; *** $p \leq 0.001$.

and Egger's tests statistics can be found in Supplementary Figs 7–10 and Supplementary Table 11, respectively.

Discussion

The present study employed a systematic review and meta-analysis framework to investigate the efficacy, tolerability and acceptability of AVATAR for AVHs. Analyses showed preliminary evidence of the efficacy of AVATAR in decreasing AVH severity at post-treatment and short-term follow-up. This was robust for the effect of bias, and after correction for potential publication bias, although the effect was moderated by control group. Notably, the analyses were performed in persons with treatment-resistant and persistent AVHs, suggesting that AVATAR may address a critical treatment gap for previously refractory symptoms. The effect size observed (Hedges' $g = -0.40$) was comparable to that of guideline-recommended psychological interventions, such as CBT (Hedges' $g = -0.34$); however, this was in populations not specifically characterised as treatment-resistant.²¹ This illustrates the compelling potential of a theory-based intervention that harnesses digital technologies in a novel and effective way to treat AVHs. Medium- and long-term follow-ups (non-significant small effect

sizes) were measured by only one study, limiting interpretability and certainty of evidence.

Analyses of effects on AVH frequency and distress showed significant small effects at post-treatment. They remained small and significant at short-term follow-up, and robust for the effect of bias (exception: distress at short-term follow-up) and potential publication bias. From a theoretical standpoint, the effects may align with potential mechanisms underlying reductions in AVH severity. The exposure to distressing avatar utterances and subsequent assertive strategies seem to have reduced both the frequency and distress of AVHs. Qualitative evidence has shown that although those affected by AVHs prioritise a decrease in the frequency of AVHs (i.e. total cessation), a decrease in AVH distress and disruption is more prioritised by service providers.⁷⁶ These findings indicate that AVATAR may support shared goals between patients and caregivers, with the potential for mutually beneficial outcomes. Of note is that the duration of AVATAR generally consisted of fewer sessions (7–12) than a typical minimal dose of 16 sessions of a psychological intervention such as CBT,⁹ emphasising the efficiency of AVATAR. Nevertheless, AVATAR is comparatively resource-intensive and may be demanding to implement at scale, because of the requirements of software, hardware and training.⁶⁴

Some concerns of publication bias were present in the AVH symptom analyses at post-treatment for AVH severity and distress. Funnel plots revealed the strong influence two small outlier studies with large effect sizes.^{65,66} This may point less toward the potential for missing studies and may instead reflect true heterogeneity of small studies. These are often able to direct more resources into treatment intensity, but are also more likely to be afflicted by methodological effects of bias.⁷⁷ Furthermore, the outlier studies had waitlist control groups, which a recent meta-analysis has shown to have the smallest within-group symptom reductions in treatment-resistant schizophrenia,⁷⁸ underlining the potential for inflated effects in these designs.

AVATAR therapy was generally well tolerated. Treatment-related adverse events and lasting symptom exacerbations were not common. However, two studies reported instances of anxiety occurring during early exposure to avatars,^{65,69} and a few cases requiring hospital admission attributable to treatment-related AVH symptom exacerbations.⁶⁹ As with other exposure-based therapies, initial increases in anxiety are not unusual and typically diminish as the treatment progresses.⁷⁹ Consistent with this, two included studies found that within-session anxiety significantly decreased over the course of the treatment.^{26,27,65} Additionally, small effect sizes favoured AVATAR in the reduction of anxiety symptoms at post-treatment and short-term follow-up. This further suggests that anxiety exacerbation is likely transient for the majority of participants. Studies also employed techniques to mitigate acute distress, such as offering a panic button or showing images of a calming beach.^{25,69} However, expanding upon these techniques to avoid overwhelming participants may be beneficial in AVATAR to minimise adverse events and attrition.

Treatment non-completion was observed in approximately a quarter of participants receiving AVATAR, which did not differ significantly in drop-out risk to control groups, supporting the general acceptability of the intervention. Notably, other exposure-based interventions report similar treatment drop-out rates (e.g. 28% for prolonged exposure in post-traumatic stress disorder⁸⁰). The aggregated study drop-out rate of 16% across intervention and control conditions closely aligns with drop-out rates of CBT in schizophrenia spectrum disorders (14%⁸¹). This suggests that AVATAR is comparable to current recommended evidence-based treatments in terms of retention.

Positive symptoms improved compared to control groups, which is likely associated with the improvement in AVH symptoms. Effect sizes were comparable to those reported for CBT in treatment-resistant schizophrenia compared with TAU and supportive counselling (Hedges' $g = -0.31$ and -0.19 , respectively⁸²). Total psychotic and negative symptoms did not show significant improvement, which may reflect the more specific focus on AVHs. In contrast, the small effect on anxiety symptoms at post-treatment and short-term follow-up makes the anxiolytic effect of exposure to distressing AVHs apparent. Significant results were not observed for social functioning and quality of life. Contrary to common impairment in schizophrenia spectrum disorders,⁸³ functional outcomes were measured in few of the included studies, reflecting a common oversight in schizophrenia spectrum disorder research.⁸⁴ Future trials should aim to consistently include functional outcome measures.

Limitations

The results of this study should be considered alongside several limitations. As a meta-analysis of under ten studies, robustness of results is not assured, requiring additional high-quality, randomised controlled trials. Likewise, publication bias tests and meta-regression tests may have been underpowered, potentially missing significant

effects.^{77,85} Another limiting factor concerns the length of measurement: only one study measured follow-ups later than 3 months post-intervention. Future studies are encouraged to perform longer follow-ups to assess the retention of effects. Although statistical heterogeneity was non-significant in the majority of analyses and often controlled by the removal of outliers, differences in treatment and follow-up durations, inclusion criteria, study designs, outcome measures and immersion were observed, potentially leading to clinical heterogeneity not assessed in these analyses. Future reviews should plan further subgroup and meta-regression analyses. Similarly, the question remains as to the comparative efficacy and tolerability of immersive virtual reality-based versus less immersive screen-based approaches, which could not be answered in this review and may be assessed with future direct comparison trials. Many included samples were small, had heterogeneous outcomes and were monocentric, with one study enrolling only 19 participants,⁶⁵ potentially introducing bias into effect sizes.⁸⁶ Future trials should be designed as large, multicentric trials with standardised outcome measures to allow for greater comparability between studies. Finally, trials were overwhelmingly conducted in Western, Educated, Industrial, Rich and Democratic (WEIRD) countries,⁸⁷ limiting the generalisability of results for all populations reporting AVHs, especially considering the strong cultural aspects apparent in AVHs.^{88,89}

In conclusion, AVATAR therapy showed efficacious and robust results for our primary outcome of the severity of AVH at post treatment, with moderate certainty of evidence. Effects were maintained into short-term follow-up, and AVH dimensions of both frequency and distress showed similar results. Efficacy profiles for other clinical and functional outcomes were mixed. These findings support the efficacy of AVATAR as a focused treatment for AVHs, although heterogeneity was apparent, and additional medium- and long-term evidence is required to assess the retention of effects.

Felix Oppner , Department of Neuroscience and Psychiatry, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; and Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany; **Sebastian Henges**, Department of Neuroscience and Psychiatry, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; and Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands; **Pawel Weinstein**, Department of Neuroscience and Psychiatry, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; and Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands; **Dana Arnheim**, Department of Neuroscience and Psychiatry, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; and Zachai Division of Psychiatry, Sheba Medical Center Israel and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; **Laura Fässler**, Department of Neuroscience and Psychiatry, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; and Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany; **Olivier Percie du Sert**, Prevention and Early Intervention Program for Psychoses, Douglas Research Centre, Montreal, Quebec, Canada; and Department of Psychiatry, McGill University, Montreal, Quebec, Canada; **Izabela Stefaniak**, Faculty of Medicine, Lazarski University, Warsaw, Poland; **Michel Sabé**, Psychiatry Department, Faculty of Medicine, University of Geneva, Geneva, Switzerland; and Division of Adult Psychiatry, Department of Psychiatry, University Hospitals of Geneva, Geneva, Switzerland; **Louise Birkedal Glenthøj**, Department of Psychology, University of Copenhagen, Copenhagen, Denmark; and VIRTU Research Group, Mental Health Center Copenhagen, Copenhagen University Hospital, Mental Health Services CPH, Copenhagen, Denmark; **Neil Thomas** , Centre for Mental Health and Brain Sciences, Swinburne University of Technology, Melbourne, Australia; **Chih-Sung Liang** , Department of Psychiatry, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan; and Department of Psychiatry, Beitou Branch, Tri-Service General Hospital, Taipei, Taiwan; **Brendon Stubbs**, Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London, London, UK; **Kerem Böge** , Department of Neuroscience and Psychiatry, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; Medical University Brandenburg-Theodor Fontane, Neuruppin, Germany; and German Center of Mental Health (DZPG), Berlin/Potsdam, Germany

Correspondence: Felix Oppner. Email: felix.opper@charite.de

First received 26 Jun 2025, final revision 17 Feb 2026, accepted 23 Feb 2026

Supplementary material

The supplementary material is available online at <https://doi.org/10.1192/bjo.2026.11014>

Data availability

Additional material is available in the online Supplementary Material. Data and scripts can be supplied upon reasonable request.

Acknowledgements

We would like to thank Professor Alex Leff for providing support and additional information on the original AVATAR trial. Additionally, we thank Dr Daniel Schulze and Dr Lars Schulze for answering early statistical questions.

Author contributions

F.O. contributed to study conceptualisation, data curation, formal analysis, methodology, project administration, visualisation and software, and wrote the original draft of the manuscript. S.H. contributed to study conceptualisation, validation and data curation, wrote the original draft of the manuscript and reviewed and edited the manuscript. P.W., L.F., M.S., N.T., C.-S.L. and B.S. reviewed and edited the manuscript. D.A. contributed to validation and reviewed and edited the manuscript. O.P.d.S., I.S. and L.B.G. contributed to data curation and reviewed and edited the manuscript. K.B. contributed to study conceptualisation, methodology, project administration, resources and supervision, and reviewed and edited the manuscript.

Funding

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Declaration of interest

L.B.G., O.P.d.S., N.T. and I.S. were directly involved with included AVATAR trials. These authors played purely advisory roles and had no influence over methodology, or rating of bias and certainty of evidence. B.S. is supported by a National Institute for Health and Care Research Advanced Fellowship (grant number NIHR301206). B.S. is on the editorial board of *Nature Exercise Science and Health*, *The Journal of Physical Activity and Health*, *Ageing Research Reviews*, *Mental Health and Physical Activity*, *The Journal of Evidence Based Medicine* and *The Brazilian Journal of Psychiatry*. B.S. has received honorarium from a co-edited book on exercise and mental illness (Elsevier), an education course and unrelated advisory work from ASICS and FitXR Ltd. L.B.G. has received honoraria from Heka-VR for clinician training in the Challenge-virtual reality-assisted therapy intervention in Denmark and internationally, is engaged in other research collaborations with Heka-VR, and has received research funding for related virtual reality-based studies. N.T. has received funding from the Australian National Health and Medical Research Council for research into AVATAR therapy. K.B. is a co-founder of the two healthcare start-ups Kiso GmbH and Mental Hub. He has also received consultancy fees and/or honoraria for lectures and/or educational materials from Boehringer Ingelheim and Angelini, as well as from publishers and training institutes for workshops, books and lectures on psychotherapy. F.O., S.H., P.W., D.A., L.F., M.S. and C.-S.L. report no competing interests.

References

- Beck AT, Rector NA. A cognitive model of hallucinations. *Cogn Ther Res* 2003; **27**: 19–52.
- Majjer K, Begemann MJH, Palmen SJMC, Leucht S, Sommer IEC. Auditory hallucinations across the lifespan: a systematic review and meta-analysis. *Psychol Med* 2018; **48**: 879–88.
- McCarthy-Jones S, Trauer T, Mackinnon A, Sims E, Thomas N, Copolov DL. A new phenomenological survey of auditory hallucinations: evidence for subtypes and implications for theory and practice. *Schizophr Bull* 2014; **40**: 231–5.
- Woods A, Jones N, Alderson-Day B, Callard F, Fernyhough C. Experiences of hearing voices: analysis of a novel phenomenological survey. *Lancet Psychiatry* 2015; **2**: 323–31.
- Larøi F, Bless JJ, Laloyaux J, Kråkvik B, Vedul-Kjelsås E, Kalhovde AM, et al. An epidemiological study on the prevalence of hallucinations in a general-population sample: effects of age and sensory modality. *Psychiatry Res* 2019; **272**: 707–14.
- Toh WL, Thomas N, Hollander Y, Rossell SL. On the phenomenology of auditory verbal hallucinations in affective and non-affective psychosis. *Psychiatry Res* 2020; **290**: 113147.
- Hasan A, Falkai P, Lehmann I. Die aktualisierte S3-Leitlinie Schizophrenie: Entwicklungsprozess und ausgewählte Empfehlungen [Revised S3 guidelines on schizophrenia: developmental process and selected recommendations]. *Nervenarzt* 2020; **91**: 26–33.
- Keepers GA, Fochtmann LJ, Anzai JM. The American Psychiatric Association Practice Guideline for the treatment of patients with schizophrenia. *Am J Psychiatry* 2020; **177**: 668–72.
- Kuipers E, Yesufu-Udechuku A, Taylor C, Kendall T. Management of psychosis and schizophrenia in adults: summary of updated NICE guidance. *BMJ* 2014; **348**: g2234.
- Samara MT, Nikolakopoulou A, Salanti G, Leucht S. How many patients with schizophrenia do not respond to antipsychotic drugs in the short term? An analysis based on individual patient data from randomized controlled trials. *Schizophr Bull* 2019; **45**: 639–46.
- Diniz E, Fonseca L, Rocha D. Treatment resistance in schizophrenia: a meta-analysis of prevalence and correlates. *Braz J Psychiatry* 2023; **45**: 448–58.
- Mørup MF, Kymes SM, Oudin Åström D. A modelling approach to estimate the prevalence of treatment-resistant schizophrenia in the United States. *PLOS One* 2020; **15**: e0234121.
- Siskind D, Orr S, Sinha S. Rates of treatment-resistant schizophrenia from first-episode cohorts: systematic review and meta-analysis. *Br J Psychiatry* 2022; **220**: 115–20.
- Bogers JPAM, Hambarian G, Walburgh Schmidt N, Vermeulen JM, Haan L. Risk factors for psychotic relapse after dose reduction or discontinuation of antipsychotics in patients with chronic schizophrenia. A meta-analysis of randomized controlled trials. *Schizophr Bull* 2023; **49**: 11–23.
- Zipursky RB, Menezes NM, Streiner DL. Risk of symptom recurrence with medication discontinuation in first-episode psychosis: a systematic review. *Schizophr Res* 2014; **152**: 408–14.
- Leucht S, Priller J, Davis JM. Antipsychotic drugs: a concise review of history, classification, indications, mechanism, efficacy, side effects, dosing, and clinical application. *Am J Psychiatry* 2024; **181**: 865–78.
- Goghari VM, Harrow M, Grossman LS, Rosen CA. 20-year multi-follow-up of hallucinations in schizophrenia, other psychotic, and mood disorders. *Psychol Med* 2013; **43**: 1151–60.
- Heilbronner U, Samara M, Leucht S, Falkai P, Schulze TG. The longitudinal course of schizophrenia across the lifespan: clinical, cognitive, and neurobiological aspects. *Harv Rev Psychiatry* 2016; **24**: 118–28.
- Nathou C, Etard O, Dollfus S. Auditory verbal hallucinations in schizophrenia: current perspectives in brain stimulation treatments. *Neuropsychiatr Treat* 2019; **15**: 2105–17.
- Bighelli I, Salanti G, Huhn M, Schneider-Thoma J, Krause M, Reitmeir C, et al. Psychological interventions to reduce positive symptoms in schizophrenia: systematic review and network meta-analysis. *World Psychiatry* 2018; **17**: 316–29.
- Turner DT, Burger S, Smit F, Valmaggia LR, Gaag M. What constitutes sufficient evidence for case formulation-driven CBT for psychosis? Cumulative meta-analysis of the effect on hallucinations and delusions. *Eur Psychiatry* 2020; **46**: 1072–85.
- Lincoln TM, Jung E, Wiesjahn M, Schlier B. What is the minimal dose of cognitive behavior therapy for psychosis? An approximation using repeated assessments over 45 sessions. *Eur Psychiatry J Assoc Eur Psychiatry* 2016; **38**: 31–9.
- Freeman D. Persecutory delusions: a cognitive perspective on understanding and treatment. *Lancet Psychiatry* 2016; **3**: 685–92.
- Smailes D, Alderson-Day B, Fernyhough C, McCarthy-Jones S, Dodgson G. Tailoring cognitive behavioral therapy to subtypes of voice-hearing. *Front Psychol* 2015; **6**: 1933.
- Leff J, Williams G, Huckvale MA, Arbutnot M, Leff AP. Computer-assisted therapy for medication-resistant auditory hallucinations: proof-of-concept study. *Br J Psychiatry* 2013; **202**: 428–33.
- Craig TKJ, Rus-Calafell M, Ward T, Leff JP, Huckvale M, Howarth E, et al. AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial. *Lancet Psychiatry* 2018; **5**: 31–40.
- Rus-Calafell M, Ward T, Zhang XC, Edwards CJ, Garety P, Craig T. The role of sense of voice presence and anxiety reduction in AVATAR therapy. *J Clin Med* 2020; **9**: 2748.
- Rus-Calafell M, Ehrbar N, Ward T, Edwards C, Huckvale M, Walke J, et al. Participants' experiences of AVATAR therapy for distressing voices: a thematic qualitative evaluation. *BMC Psychiatry* 2022; **22**: 356.
- Tsang A, Bucci S, Branitsky A. The relationship between appraisals of voices (auditory verbal hallucinations) and distress in voice-hearers with schizophrenia-spectrum diagnoses: a meta-analytic review. *Schizophr Res* 2021; **230**: 38–47.
- Spark J, Pot-Kolder R, Dzafic I, Nelson B, Byrne LK, Lum JAG. Virtual reality for the treatment of positive symptoms of psychosis: a meta-analysis of trials. *Curr Treat Options Psychiatry* 2025; **12**: 10.

- 31 Zeka F, Clemmensen L, Valmaggia L, Veling W, Hjorthøj C, Glenthøj LB. The effectiveness of immersive virtual reality-based treatment for mental disorders: a systematic review with meta-analysis. *Acta Psychiatr Scand* 2025; **151**: 210–30.
- 32 Aali G, Kariotis T, Shokraneh F. AVATAR therapy for people with schizophrenia or related disorders. *Cochrane Database Syst Rev* 2020; **5**: CD011898.
- 33 Hsu T-W, Tseng P-T, Hsu C-W, Yang F-C, Changchien T-C, Lin Y-H, et al. AVATAR therapy for medication-resistant auditory hallucination in patients with psychosis: a systematic review and meta-analysis. *Schizophrenia* 2025; **12**: 1.
- 34 Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA. 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; **372**: n71.
- 35 Bramer WM, De Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. *J Med Libr Assoc* 2018; **106**: 531–41.
- 36 Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan – a web and mobile app for systematic reviews. *Syst Rev* 2016; **5**: 210.
- 37 Polanin JR, Pigott TD, Espelage DL, Grotzinger JK. Best practice guidelines for abstract screening large-evidence systematic reviews and meta-analyses. *Res Synth Methods* 2019; **10**: 330.
- 38 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; **20**: 37–46.
- 39 Maldonado AD, Marzo PF, Andrés AM. DeltaMAN: Delta Measurement of Agreement for Nominal Data. CRAN R-Project, 2022 (<https://cran.r-project.org/web/packages/DeltaMAN/index.html>)
- 40 Haddock G, McCarron J, Tarrier N, Faragher EB. Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychol Med* 1999; **29**: 879–89.
- 41 Guaiana G, Abbatecola M, Aali G, Tarantino F, Ebuonyi ID, Lucarini V, et al. Cognitive behavioural therapy (group) for schizophrenia. *Cochrane Database Syst Rev* 2022; **7**: CD009608.
- 42 Mayer SF, Corcoran C, Kennedy L, Leucht S, Bighelli I. Cognitive behavioural therapy added to standard care for first-episode and recent-onset psychosis. *Cochrane Database Syst Rev* 2024; **3**: CD015331.
- 43 Cooper AA, Conklin LR. Dropout from individual psychotherapy for major depression: a meta-analysis of randomized clinical trials. *Clin Psychol Rev* 2015; **40**: 57–65.
- 44 Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011; **342**: d549.
- 45 Hedges LV. Distribution theory for glass's estimator of effect size and related estimators. *J Educ Stat* 1981; **6**: 107–28.
- 46 Cohen J. *Statistical Power Analysis for the Behavioral Sciences* 2nd ed. Routledge, 1988.
- 47 Lin L, Chu H. Meta-analysis of proportions using generalized linear mixed models. *Epidemiology* 2020; **31**: 713–7.
- 48 Weber F, Knapp G, Ickstadt K, Kundt G, Glass A. Zero-cell corrections in random-effects meta-analyses. *Res Synth Methods* 2020; **11**: 913–9.
- 49 Higgins JP, Eldridge S, Li T. Including variants on randomized trials. In *Cochrane Handbook for Systematic Reviews of Interventions* 1st ed. (eds JPT Higgins, J Thomas, J Chandler, M Cumpston, T Li, MJ Page, et al.): 569–93. Wiley, 2019.
- 50 Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954; **10**: 101–29.
- 51 Higgins JPT. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557–60.
- 52 Posit Team. *RStudio: Integrated Development Environment for R*. Posit Team, 2025 (<https://posit.co/products/open-source/rstudio/?sid=1>).
- 53 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010; **36**: 1–48.
- 54 Wickham H, Averick M, Bryan J. Welcome to the tidyverse. *J Open Source Softw* 2019; **4**: 1686.
- 55 Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. *Res Synth Methods* 2010; **1**: 112–25.
- 56 Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; **366**: l4898.
- 57 McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods* 2021; **12**: 55–61.
- 58 Dickersin K, Min Y-I. Publication bias: the problem that won't go away. *Ann N Y Acad Sci* 1993; **703**: 135–48.
- 59 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629–34.
- 60 Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; **56**: 455–63.
- 61 Guyatt GH, Oxman AD, Vist GE. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; **336**: 924–6.
- 62 McMaster University, Evidence Prime. *GRADEpro Guideline Development Tool*. Evidence Prime, 2025 (<https://www.gradepr.org>).
- 63 McHugh ML. Interrater reliability: the kappa statistic. *Biochem Medica* 2012; **22**: 276.
- 64 Garety PA, Edwards CJ, Jafari H. Digital AVATAR therapy for distressing voices in psychosis: the phase 2/3 AVATAR2 trial. *Nat Med* 2024; **30**: 3658–68.
- 65 Percie du Sert O, Potvin S, Lipp O, Dellazizzo L, Laurelli M, Breton R, et al. Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: a pilot clinical trial. *Schizophr Res* 2018; **197**: 176–81.
- 66 Stefaniak I, Sorokosz K, Janicki A, Wciórka J. Therapy based on avatar-therapist synergy for patients with chronic auditory hallucinations: a pilot study. *Schizophr Res* 2019; **211**: 115–7.
- 67 Dellazizzo L, Potvin S, Phraxayavong K, Dumais A. One-year randomized trial comparing virtual reality-assisted therapy to cognitive-behavioral therapy for patients with treatment-resistant schizophrenia. *NPJ Schizophr* 2021; **7**: 9.
- 68 Liang N, Li X, Guo X. Visual P300 as a neurophysiological correlate of symptomatic improvement by a virtual reality-based computer AT system in patients with auditory verbal hallucinations: a pilot study. *J Psychiatr Res* 2022; **151**: 261–71.
- 69 Smith LC, Vernal DL, Mariegaard LS, Christensen AG, Jansen JE, Schytte G, et al. Immersive virtual reality-assisted therapy targeting persistent auditory verbal hallucinations in patients diagnosed with schizophrenia spectrum disorders in Denmark: the challenge assessor masked, randomized clinical trial. *Lancet Psychiatry* 2025; **12**: 557–67.
- 70 Burgess-Barr S, Nicholas E, Venus B, Singh N, Nethercott A, Taylor G, 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**: 8.
- 71 Correll CU, Arango C, Fagioli A, Giordano GM, Leucht S, Salazar de Pablo G. Finding the right setting for the right treatment during the acute treatment of individuals with schizophrenia: a narrative review and clinical practice guideline. *Neuropsychiatr Dis Treat* 2024; **20**: 1293–307.
- 72 McDonagh MS, Dana T, Selph S, Devine EB, Cantor A, Bougatsos C, et al. *Treatments for Schizophrenia in Adults: A Systematic Review*. Agency for Healthcare Research and Quality, 2017 (<https://www.ncbi.nlm.nih.gov/books/NBK487628/>).
- 73 The GRADE Working Group. *GRADE Handbook for Grading Quality of Evidence and Strength of Recommendations*. The GRADE Working Group, 2013 (<https://gdt.gradepr.org/app/handbook/handbook.html>).
- 74 Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis* 1st ed. Wiley, 2009.
- 75 Smith DA, Mar CM, Turoff BK. The structure of schizophrenic symptoms: a meta-analytic confirmatory factor analysis. *Schizophr Res* 1998; **31**: 57–70.
- 76 Longden E, Branitsky A, Sheaves B, Chauhan N, Morrison AP. Preferred treatment outcomes in psychological therapy for voices: a comparison of staff and service-user perspectives. *Psychosis* 2024; **16**: 107–17.
- 77 Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011; **343**: d4002.
- 78 Schütz A, Salahuddin NH, Priller J, Bighelli I, Leucht S. The role of control groups in non-pharmacological randomised controlled trials of treatment-resistant schizophrenia: a systematic review and meta-analysis. *Psychiatry Res* 2024; **339**: 116069.
- 79 Heing I, Knappe S, Hoyer J. Effective – and tolerable: acceptance and side effects of intensified exposure for anxiety disorders. *Behav Ther* 2023; **54**: 427–43.
- 80 Varker T, Jones KA, Arjmand H-A, Hinton M, Hiles SA, Freijah I, et al. Dropout from guideline-recommended psychological treatments for posttraumatic stress disorder: a systematic review and meta-analysis. *J Affect Disord Rep* 2021; **4**: 100093.
- 81 Cuijpers P, Harrer M, Miguel C, Ciharova M, Papola D, Basic D, et al. Cognitive behavior therapy for mental disorders in adults: a unified series of meta-analyses. *JAMA Psychiatry* 2025; **82**: 563.
- 82 Salahuddin NH, Schütz A, Pitschel-Walz G. Psychological and psychosocial interventions for treatment-resistant schizophrenia: a systematic review and network meta-analysis. *Lancet Psychiatry* 2024; **11**: 545–53.
- 83 Laws KR, Darlington N, Kondel TK, McKenna PJ, Jauhar S. Cognitive behavioural therapy for schizophrenia – outcomes for functioning, distress and quality of life: a meta-analysis. *BMC Psychol* 2018; **6**: 32.

- 84 Bighelli I, Wallis S, Reitmeir C, Schwermann F, Salahuddin NH, Leucht S. Effects of psychological treatments on functioning in people with schizophrenia: a systematic review and meta-analysis of randomized controlled trials. *Eur Arch Psychiatry Clin Neurosci* 2023; **273**: 779–810.
- 85 Deeks JJ, Higgins JP, Altman DG. Analysing data and undertaking meta-analyses. In *Cochrane Handbook for Systematic Reviews of Interventions* 1st ed. (eds JPT Higgins, J Thomas, J Chandler, M Cumpston, T Li, MJ Page, et al): Chapter 10. Wiley, 2019.
- 86 Lin L. Bias caused by sampling error in meta-analysis with small sample sizes. *PLOS One* 2018; **13**: e0204056.
- 87 Muthukrishna M, Bell AV, Henrich J. Beyond western, educated, industrial, rich, and democratic (WEIRD) psychology: measuring and mapping scales of cultural and psychological distance. *Psychol Sci* 2020; **31**: 678–701.
- 88 Khaled SM, Brederoo SG, Yehya A, Alabdulla M, Woodruff PW, Sommer IEC. Cross-cultural differences in hallucinations: a comparison between middle eastern and European community-based samples. *Schizophr Bull* 2023; **49**: S13.
- 89 Larøi F, Luhrmann TM, Bell V, Christian WA Jr, Deshpande S, Fernyhough C, et al. Culture and hallucinations: overview and future directions. *Schizophr Bull* 2014; **40**: S213–20.