Schizophrenia Bulletin vol. 46 no. 5 pp. 1072–1085, 2020
doi:10.1093/schbul/sbaa045
Advance Access publication 27 March 2020

What Constitutes Sufficient Evidence for Case Formulation–Driven CBT for Psychosis? Cumulative Meta-analysis of the Effect on Hallucinations and Delusions

David T. Turner*[,1], Simone Burger[1], Filip Smit[1–3], Lucia R. Valmaggia[4], and Mark van der Gaag[1,5]

[1]Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [2]Department of Epidemiology and Biostatistics, Amsterdam Public Health Research Institute, Amsterdam Medical Centers, Amsterdam, The Netherlands; [3]Trimbos Institute, Netherlands Institute of Mental Health, Utrecht, The Netherlands; [4]Department of Psychology, King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, UK; [5]Parnassia Psychiatric Institute, The Hague, The Netherlands

*To whom correspondence should be addressed; Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands; tel: +31-20-5985950, fax: +31-20-5988758; e-mail: d.t.turner@vu.nl

Objective: **Following 2 decades of research on cognitive behavioral therapy for psychosis (CBTp), it is relevant to consider at which point the evidence base is considered *sufficient*. We completed a cumulative meta-analysis to assess the sufficiency and stability of the evidence base for hallucinations and delusions.*Method:* We updated the systematic search from our previous meta-analytic review from August 2013 until December 2019. We identified 20 new randomized controlled trials (RCTs) resulting in inclusion of 35 RCTs comparing CBTp with treatment-as-usual (TAU) or active controls (AC). We analyzed data from participants with psychosis ($N = 2407$) over 75 conventional meta-analytic comparisons. We completed cumulative meta-analyses (including fail-safe ratios) for key comparisons. Publication bias, heterogeneity, and risk of bias were examined.*Results:* Cumulative meta-analyses demonstrated sufficiency and stability of evidence for hallucinations and delusions. The fail-safe ratio demonstrated that the evidence base was sufficient in 2016 for hallucinations and 2015 for delusions. In conventional meta-analyses, CBTp was superior for hallucinations ($g = 0.34$, $P < .01$) and delusions ($g = 0.37$, $P < .01$) when compared with any control. Compared with TAU, CBTp demonstrated superiority for hallucinations ($g = 0.34$, $P < .01$) and delusions ($g = 0.37$, $P < .01$). Compared with AC, CBT was superior for hallucinations ($g = 0.34$, $P < .01$), but not for delusions although this comparison was underpowered. Sensitivity analyses for case formulation, primary outcome focus, and risk of bias demonstrated increases in effect magnitude for hallucinations.*Conclusions:* The evidence base for the effect of CBTp on hallucinations and delusions demonstrates sufficiency and stability across comparisons, suggesting limited value of new trials evaluating generic CBTp.**

Key words: schizophrenia/randomized controlled trials/psychological intervention/positive symptoms/systematic review

Introduction

It is now approximately 20 years since the evidence base for cognitive behavioral therapy for psychosis (CBTp) began to accumulate and, as randomized controlled trials (RCTs) continue to proliferate, it is relevant to consider at which point the evidence base is considered *sufficient*. Our previous meta-analytic review demonstrated the efficacy of individually tailored, case formulation–based CBTp in reducing hallucinations (Hedge's $g = 0.44$, $P < .005$) and delusions ($g = 0.36$, $P < .05$) when RCTs were focused on specific symptom reduction.[1] These findings were broadly in line with existing meta-analytic results for positive symptoms.[2,3] We concluded that CBTp was an efficacious intervention for hallucinations and delusions, although the lower magnitude of effect for delusions and the absence of a significant effect compared with active treatments led us to conclude that delusions may be less amenable to change via CBTp than hallucinations.

Roughly, 6 years have elapsed since our previous review. During this time, a number of new RCTs have been published in this research field. These include

© The Author(s) 2020. Published by Oxford University Press on behalf of the Maryland Psychiatric Research Center.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

trials employing the typical implementation of individually case-formulated CBTp in Western mental health care systems as were prevalent in the former review alongside a range of trials in new settings and/or employing new styles of intervention, eg, culturally adapted CBTp in Pakistan[4] or virtual-reality-based CBTp.[5] There remains well-documented controversy[6] over the effectiveness and implementation of CBTp; both the UK National Institute for Health and Care Excellence[7] and the British Psychological Society *Understanding Psychosis and Schizophrenia* report[8] recommend CBTp, whereas the Cochrane Collaboration maintain that meta-analytic results are neither clear nor robust enough to recommend CBTp over standard care.[9] Recent literature addressing this controversy argues the importance of attending to methodological issues including blinding, inclusion criteria and prespecification of methods.[6]

Cumulative meta-analysis is a technique allowing estimation of both the sufficiency and stability of meta-analytic evidence. This technique was first notably applied to treatment trials for myocardial infarction.[10] The method has since been applied as a means of statistically estimating the point at which there is sufficient evidence to conclude that an intervention is efficacious while also estimating the stability of the effect size over time.[11,12] In light of the further accumulation of trials, we concluded that the application of cumulative meta-analysis to the CBTp field is warranted.

We first aimed to update our 2014 review to assimilate the new body of research and therefore provide an up-to-date estimation of the impact of CBTp upon hallucinations and delusions. We also employed cumulative meta-analysis to comment on the sufficiency of the existing evidence base in demonstrating efficacy and the stability of the evidence over time. A secondary objective was to provide a range of sensitivity analyses to allow more specific estimation of effects under prespecified conditions such as individually tailored case formulation, primary outcome focus, blinded RCTs, and RCTs with minimal risk of bias.

Methods

We provide a systematic review including both conventional and cumulative meta-analyses based on PRISMA guidelines.[13] A protocol for this review was registered at the Open Science Framework (https://osf.io/nwxbz/).

Systematic Search

Our previous meta-analytic review in this area completed a systematic search on August 3, 2013.[1] We repeated the systematic search from this date until December 11, 2019 across the same 3 databases included in 2013 (PubMed, Embase, and PsychInfo). We considered reference lists of published reviews alongside our accumulation of newly published trials via automatic update notifications and expert knowledge via professional networks. We entered a relevant range of text variations of the following key search terms via while utilizing Boolean operators, MeSH terms, and exploded terms and limit setting based on specific options within each database: (1) cognitive behavioral therapy, (2) auditory hallucinations OR delusions, and (3) RCTs. Exemplary search strings are included in supplementary materials.

Inclusion/Exclusion Criteria

We included (1) RCTs comparing (2) cognitive behavioral therapy with (3) treatment-as-usual (TAU) or an active control condition (eg, supportive counseling or psychoeducation) for (4) patients diagnosed with schizophrenia-spectrum disorders which (5) assessed hallucinations and/or delusions as post-treatment outcome. Schizophrenia-spectrum disorders included schizophrenia, schizoaffective disorder, delusional disorder, brief psychotic disorder or psychosis not otherwise specified. We included only studies published in peer-reviewed journals. Conference abstracts were excluded. We also excluded trials that (1) focused on a primary diagnosis of alcohol or substance use dependency; (2) included ultra-high risk patients or focused on prevention of psychosis; and (3) replaced the core of CBT (ie, identifying and challenging of maladaptive beliefs) with alternative psychological interventions, eg, social skill training or mindfulness. We utilized the definition of CBTp applied in our previous meta-analytic research.[1]

Study Selection

The PRISMA diagram (figure 1) depicts the study selection process. Two authors (D.T. and M.v.d.G.) utilized the Rayyan (rayyan.qcri.org) web application to facilitate the study selection process. Abstracts were first screened for duplicates then relevance before a sample of full text PDFs were checked against the inclusion and exclusion criteria. Conflicts in inclusion were resolved via discussion. We attempted to contact the authors of one RCT due to PSYRATS subscales being unavailable in the manuscript but received no response.[14]

Data Extraction

Two authors (D.T. and S.B.) independently completed the data extraction for new trials included since 2013. The data from trials included in the 2013 review were also checked for consistency by both authors, and any inconsistencies were investigated and corrected. Spreadsheets utilized in the previous meta-analyses were adapted and updated for use in the current review. We contacted one author for unavailable data although on closer inspection

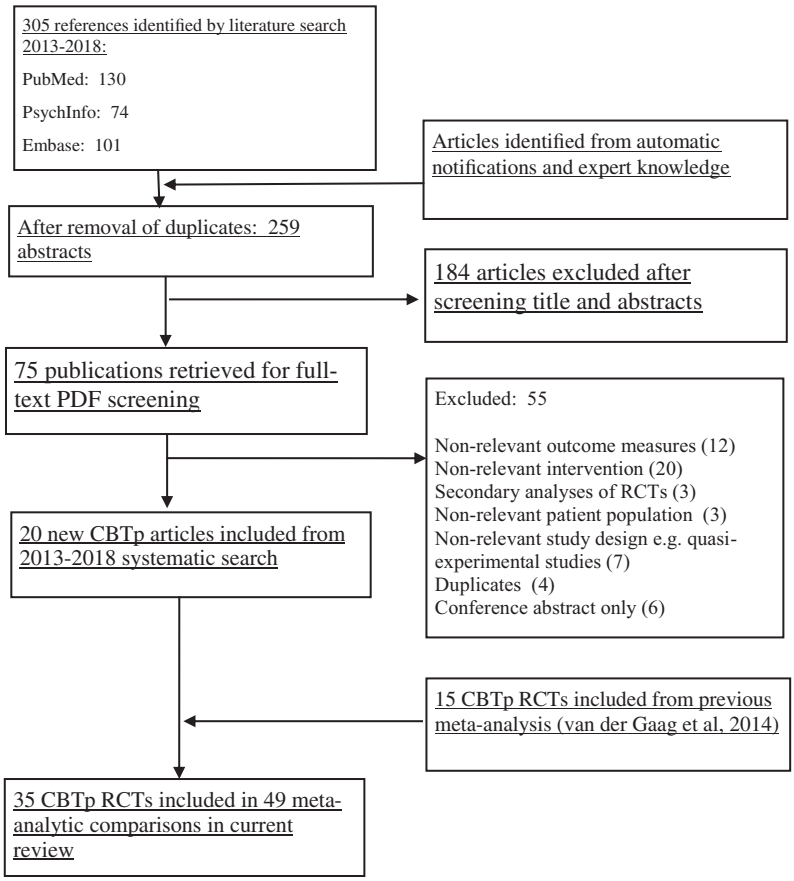

Fig. 1. Flowchart of inclusion of studies.

of the manuscript, the intervention in this trial did not meet inclusion criteria. Data were extracted on study characteristics (year of publication, country, sample characteristics, format [individual or group], duration, application of case formulation, primary vs. secondary focus and intervention style) and post-treatment outcome data.

Outcome Measures

Although a considerable proportion of meta-analytic research on cognitive behavioral therapy for psychosis (CBTp) has focused on its effect in reducing the positive or negative symptoms of psychosis,[2,3,15] there has been less focus on the more specific, discrete outcomes of hallucinations and delusions. It has been suggested that diagnostically based tools such as the Positive and Negative Syndromes Scale (PANSS)[16] provide less comprehensive measurement of psychotic symptomatology than symptom-specific outcome measures such as the Psychotic Symptoms Rating Scales (PSYRATS).[17,18] Our primary outcomes were therefore hallucinations and delusions. We extracted all outcome measures that reported hallucinations or delusions as independent scales or subscales. We did not include outcome measures that subsumed items on hallucinations or delusions in broader subcategories such as positive symptoms (eg, the PANSS). In instances where 2 hallucinations or 2 delusions scales were reported, data from both were extracted, and an average pooled effect size was calculated. All scales included were continuous outcomes.

Risk of Bias Assessment

To account for risk of bias among the included RCTs, we applied an adapted version of the Cochrane Risk of Bias tool. The final 2 items of the tool (selective outcome reporting and other sources of bias) were omitted due to limited evidence regarding their impact on validity for meta-analytic comparisons.[19] Utilization of the 4 key areas of bias (namely sequence generation, allocation concealment, blinding of assessors, and incomplete outcome data) provided the opportunity for clear sensitivity analyses as applied in previous meta-analytic reviews.[2,20] Risk of bias was assessed independently by 2 authors (D.T. and M.v.d.G.). Conflicts were resolved via discussion. Risk of bias items were rated *low risk* (0) or *high risk* (1), contributing to a total score of 0–4 for each RCT. Items that were unclear in the published manuscripts were rated conservatively as high risk. Due to an alternative method of risk of bias assessment being employed in the earlier review, risk of bias was assessed over the whole sample of RCTs.

Meta-analyses

Our strategy for analysis was to move gradually from inclusive comparisons to more exclusive sensitivity analyses for a number of criteria based on (1) relevant study characteristics and (2) risk of bias. These sensitivity analyses were designed to provide information relevant to our aforementioned research objectives, namely our focus on individually tailored, case formulation–driven CBT with hallucinations and delusions as primary outcome. We therefore first analyzed all eligible RCTs each for hallucinations and delusions. We then completed sensitivity analyses examining TAU only, active controls only, case formulation only, and primary outcomes only. When study availability allowed sufficient number of RCTs for comparison, we also included smaller categories including group CBT only, secondary outcomes only, self-help CBT only, and virtual-reality CBT (VR-CBT) only. We note that the minimum number of RCTs required for adequate meta-analytic comparisons is suggested as approximately 5.[21] Comparisons we reported in this section, which fell below this 5 RCTs, were therefore provided only for indicative information regarding current best estimates. When possible, based on RCT availability, we also performed sensitivity analyses including only RCTs with low risk of bias (one of more items scored on the risk of bias tool) and no known risk of bias (no items scored on the risk of bias tool).

All meta-analytic comparisons were completed using the Comprehensive Meta-analysis (CMA) version 3.3.070 computer software package. CMA provides an aggregated effect size estimating the pooled mean difference between treatment and control groups at post-treatment using Hedge's *g*, which is an estimate of the standardized mean difference between study groups. Hedge's *g* is recognized as providing a more accurate effect estimation in small samples than alternative methods for continuous measures such as Cohen's *d*. We utilized the .05 alpha level for all comparisons with 95% confidence intervals (CIs) provided. We also employed a random-effects model in all comparisons due to the expectation of between-study variance.

Cumulative Meta-analysis

To assess the sufficiency and stability of meta-analytic evidence for CBTp for hallucinations and delusions, we completed cumulative meta-analyses for each outcome. The cumulative meta-analysis function in *CMA* was utilized. RCTs were listed by year of publication, and a pooled effect size in Hedge's *g* was calculated for the point at which each new study was chronologically added to the evidence base. The cumulative forest plots (supplementary figures 5–9) provide a visual representation of the stability of evidence as the RCT evidence based has accumulated. We also followed Muellerleile and Mullen's[11] recommendations for calculating the fail-safe ratio from Rosenthal's[22] standard to estimate the sufficiency of the evidence base as each RCT was added. The fail-safe ratio provides an estimate of sufficiency of evidence when the ratio surpasses a value of 1.0.

Publication Bias, Heterogeneity, and Power

We utilized the *Q* statistic and the I^2 statistic to assess heterogeneity. We examined publication bias to estimate the potential impact of unpublished RCTs. We applied power calculations to estimate the number of RCTs required for adequate power in each comparison.[23] Full information for these procedures is provided in supplementary materials.

Results

Study Selection

After the automated removal of duplicates, the updated search resulted in 305 new citations being retrieved for abstract screening, of which 184 were excluded and 75 full-text PDFs were retrieved. Following careful matching of exclusion and inclusion criteria, 20 new studies published since the previous meta-analysis were included meaning a total of 35 RCTs were included in this meta-analytic review. The total amount of participants measured at post-treatment was $N = 2407$, which included 1205 patients who received CBT and 1202 patients who received TAU or an active control. We analyzed their data over 75 meta-analytic comparisons.

Selected study characteristics are provided in table 1. Twenty-eight RCTs (80%) applied individually tailored case formulation, whereas 9 studies (26%) did not. Thirty-one RCTs (89%) targeted hallucinations and/or delusions as primary outcome, whereas 6 (17%) targeted these outcomes as secondary. Twenty-nine RCTs utilized TAU as the comparison condition, 6 RCTs compared CBT with active controls or other psychological interventions, whereas 2 RCTs included both. Active control treatments included supportive counseling,[24–29] psychoeducation,[30] befriending,[17] and virtual-reality exposure.[31] Only one RCT explicitly excluded participants taking antipsychotic medication from the CBTp treatment group,[32] therefore indicating that CBTp was broadly provided as adjunctive to standard care.

Risk of bias varied among RCTs although the majority (24 RCTs; 67%) achieved the best possible risk of bias score on the adapted Cochrane tool. A further 8 RCTs (23%) scored one risk of bias item, whereas 2 RCTs (6%) scored 2 items, 2 RCTs (6%) scored 3 items, and 1 RCT (3%) scored 4 items, indicating the highest possible score on the tool. Risk of bias assessment scores are provided in tabular form in supplementary materials.

Effect of CBT on Hallucinations

Table 2 provides an overview of all meta-analytic results of the effect of CBTp on hallucinations. Supplementary

Table 1. Selected Study Characteristics of CBTp RCTs for Hallucinations and Delusions

Author	Year	Format	Duration Intervention	Experimental Condition				Control Condition				Country	CF	Bias Risk 0–4	Selected Outcome Measure
				CBT Format	CBTp	Age	Male Sex	Control Format	Control	Age	Male Sex				
					N	Mean (SD)	%		*N*	Mean (SD)	%				
Lewis et al [29]	2002	Indiv	15–20 h in 5 wk	CBT	101	29.1	71%	[1] SC [2] TAU	[1] 106 [2] 102	[1] 27.2 [2] 27.0	[1] 71% [2] 68%	UK	Y	0	PSYRATS
Durham et al[24]	2003	Indiv	9 mo	CBT	22	36.0 (10.0)	68%	[1] SC [2] TAU	[1] 23 [2] 21	[1] 37.0 (11.2) [2] 36.0 (10.2)	[1] 65% [2] 71%	UK	Y	1	PSYRATS
Trower et al[33]	2004	Indiv	6 mo	CT CH	18	36.6 (10.3)	56%	TAU	20	35.1 (10.4)	70%	UK	Y	1	PSYRATS
Cather et al[30]	2005	Indiv	16 weekly sessions	fCBT	15	40.4 (12.0)	57%	PE	13	40.4 (12.0)	57%	USA	Y	1	PSYRATS
Wykes et al[34]	2005	Group	10 wk	CBT	45	39.7 (10.8)	53%	TAU	40	39.7 (10.1)	65%	UK	N		PSYRATS
Valmaggia et al[25]	2005	Indiv	6 mo	CBT	35	35.5 (10.8)	77%	SC	23	35.5 (11.4)	61%	NL	Y	0	PSYRATS
McLeod et al[35]	2007	Group	8 weekly sessions	CBT	10	n.a.	n.a.	TAU	10	n.a.	n.a.	UK	N	3	PSYRATS
O'Connor et al[26]	2007	Indiv	24 weekly sessions	CBT	12	40 (9.4)	45%	SC	12	36.8 (13.5)	67%	CAN	Y	3	MADS
Garety et al[36]	2008	Indiv	20 in 9 mo	CBT	60 H 85 D	39.1 (10.3)	71%	TAU	60 H 85 D	37.1 (10.9)	72%	UK	Y	0	PSYRATS
Penn et al[27]	2009	Group	12 wk	CBT	32	41.7 (11.8)	53%	SC	33	39.6 (15.7)	49%	USA	N	0	PSYRATS
Haddock et al[17]	2009	Indiv	17 sessions	CBT	38	35.7 (12.5)	86%	SC	39	33.9 (9.7)	86%	UK	Y	0	PSYRATS
Peters et al[27]	2010	Indiv	6 mo	CBT	36	34.0 (9.8)	72%	TAU	38	39.6 (10.2)	53%	UK	Y	2	BAVQ-R
Foster et al[37]	2010	Indiv	4 wk	CBT	9	40.0 (10.0)	58%	TAU	11	39.1 (9.2)	58%	UK	N	2	PSYRATS
Lincoln et al[38]	2012	Indiv	29 sessions	CBT	40	33.2 (10.4)	55%	TAU	40	33.1 (10.9)	58%	GER	Y	0	PDI
Krakvik et al[39]	2013	Indiv	6 mo 20 sessions	CBT	23	35.3 (8.9)	65%	TAU	22	37.5 (11.2)	64%	NOR	Y	1	PSYRATS
Rathod et al[40]	2013	Indiv	16 weekly sessions	CA-CBT	17	31.4 (12.4)	63%	TAU	18	35.6 (10.7)	59%	UK	Y	0	CPRS DHS
Leff et al[41]	2013	Indiv	6 weekly sessions	Avatar CT	14	n.a.	n.a	TAU	12	n.a.	n.a	UK	Y	1	PSYRATS
Morrison et al[32]	2014	Indiv	9 mo	CBT	37	33.0 (13.1)	46%	TAU	37	29.7 (12.0)	59%	UK	Y	0	PSYRATS
Birchwood et al[42]	2014	Indiv	9 mo	CT CH	98	38.8 (12.2)	62%	TAU	99	35.9 (11.9)	53%	UK	Y	0	PSYRATS
Freeman et al[43]	2014	Indiv	6 sessions	CBT confidence	15	41.9 (11.5)	73%	TAU	15	41.5 (13.1)	60%	UK	Y	0	PSYRATS
Tarrier et al[44]	2014	Indiv	24 sessions	CBT suicide	25	32.6 (11.7)	n.a.	TAU	24	37.3 (14.2)	n.a.	UK	Y	0	PSYRATS
Freeman et al[45]	2015	Indiv	8 sessions	CBT sleep	24	39·6 (11.6)	67%	TAU	26	42·2 (13.5)	69%	UK	N	0	PSYRATS
Naeem et al[46]	2015	Indiv	16 weekly sessions	CBT	53	42.0 (11.6)	17%	TAU	49	38.6 (12.0)	13%	CAN	Y	0	PSYRATS
Habib et al[47]	2015	Indiv	10–16 sessions	CBT	21	33.5 (10.5)	44%	TAU	21	30.2 (6.7)	56%	PAK	Y	0	PSYRATS
Freeman et al[45]	2015	Indiv	8 weekly sessions	CBT-W	73	40.9 (10.5)	58%	TAU	77	42.1 (12.2)	57%	UK	Y	0	GPTS and PSYRATS
Waller et al[48]	2015	Indiv	6 wk	CBT-TW	20	39.1 (10.5)	75%	TAU	11	43.0 (10.7)	64%	UK	Y	0	DC, DD, and DP

Table 1. Continued

Author	Year	Format	Duration Intervention	Experimental Condition				Control Condition				Country	CF	Bias Risk 0–4	Selected Outcome Measure
				CBT Format	CBTp	Age	Male Sex	Control Format	Control	Age	Male Sex				
					N	Mean (SD)	%		*N*	Mean (SD)	%				
Naeem et al[49]	2016	Indiv	12–16 sessions	CBT-GSH	18	42.0 (11.5)	44%	TAU	15	38.6 (12.0)	60%	CAN	Y	0	PSYRATS
Freeman et al[31]	2016	Indiv	1 session	VR-CBT	15	42.1 (13.4)	67%	Exposure	15	40.6 (14.4)	67%	UK	Y	0	PSYRATS
Hayward et al[50]	2017	Indiv	16 weekly sessions	Relating therapy	14	41 (n.p)	43%	TAU	15	43 (n.p.)	67%	UK	Y	0	PSYRATS
Hazell et al[51]	2017	Indiv	8 sessions	CBT-GSH	14	39.1 (10.2)	29%	WL	14	45.9 (13.5)	50%	UK	N	0	HPSVQ
Gottlieb et al [52]	2017	Indiv	10 skill modules	eCBT	19	43.8 (13.2)	47%	TAU	18	40.3 (11.7)	78%	USA	N	1	PSYRATS
Pot-Kolder et al[5]	2018	Indiv	16 sessions	VR-CBT	58	36.5 (10)	69%	TAU	58	39.5 (10)	72%	NL	Y	0	ESM
Morrison et al[53]	2018	Indiv	9 mo	CBT	242	42.2 (10.7)	73%	TAU	245	42.8 (10.4)	71%	UK	Y	0	PSYRATS
Husain et al[4]	2017	Indiv	12 weekly sessions	CBT	18	34.1 (9.55)	78%	TAU	18	30.5 (8.15)	55.6%	PAK	Y	0	PSYRATS
Craig et al[28]	2018	Indiv	12 weekly sessions	Avatar CT	75	42.5 (10.7)	76%	SC	75	42.9 (11.2)	60%	UK	Y	0	PSYRATS
Wong et al[54]	2019	Group	7 weekly session + booster	CBT	25	30.6 (10.6)	24%	PE	23	35.1 (12.9)	48%	HK	N	1	PSYRATS, BAVQ

Note: BAVQ-R, Beliefs About Voices Questionnaire-Revised; CA-CBT, culturally adapted CBT; CAN, Canada; CBT, cognitive behavioral therapy; CBT-GSH, cognitive-behavioral guided self-help; CBT-I, cognitive behavioral therapy for insomnia; CBT-TW, "Thinking Well" cognitive-behavioral therapy; CBT-W, cognitive-behavioral therapy for worry; CH, command hallucinations; CPRS, Comprehensive Psychopathological Rating Scale; CT, cognitive therapy; D, delusions; DC, delusional conviction; DD, delusional distress; DHS, delusions and hallucinations scale; DP, delusional preoccupation; eCBT, online cognitive-behavioral therapy; ESM, experience sampling method; fCBT, functional cognitive behavioral therapy; GER, Germany; GPTS, Green et al Paranoid Thoughts Scale; H, hallucinations; HK, Hong Kong; HPSVQ, Hamilton Program for Schizophrenia Voices Questionnaire; KOR, SK, South Korea; MADS, Maudsley Assessment of Delusions Scale; n.a., not applicable; n.p., not provided; NL, Netherlands; NOR, Norway; PAK, Pakistan; PDI, Peters et al Delusion Inventory; PE, psychoeducation; PSYRATS, Psychotic Symptoms Rating Scales; SC, supportive counseling; TAU, treatment-as-usual; VR-CBT, virtual-reality-based cognitive behavioral therapy; WL, waiting list. BAVQ comparisons included only resistance, omnipotence, and malevolence subscales.

Table 2. Effect Sizes of CBTp for Auditory Hallucinations

	N	*g*	95% CI	*Z*	*Q*-Value	I^2 (%)
Main comparison with all eligible RCTs						
Any risk of bias score included	28	0.34**	0.20, 0.49	4.63	52.83**	49
High risk of bias (>1)[a]	26	0.34**	0.19, 0.49	4.43	51.12**	51
Lowest risk of bias (0)[b]	19	0.40**	0.22, 0.58	4.40	41.49**	59
CBTp vs TAU						
Any risk of bias score included	22	0.35**	0.18, 0.52	4.00	45.94**	54
High risk of bias (>1)[a]	20	0.34**	0.17, 0.52	3.77	44.24**	58
Lowest risk of bias (0)[b]	14	0.41**	0.19, 0.63	3.65	36.90**	65
CBTp vs active intervention						
Any risk of bias score included	8	0.34**	0.15, 0.53	3.58	7.03	0
High risk of bias (>1)[a]	8	0.34**	0.15, 0.53	3.58	7.03	0
Lowest risk of bias (0)[b]	5	0.42**	0.20, 0.64	3.70	4.15	4
CBTp with hallucinations as primary outcome[c]						
Any risk of bias score included	23	0.40**	0.24, 0.56	4.90	40.42*	46
High risk of bias (>1)[a]	21	0.40**	0.23, 0.57	4.66	38.84**	49
Lowest risk of bias (0)[b]	14	0.51**	0.32, 0.70	5.22	25.67*	49
CBTp with individualized case formulation[d]						
Any risk of bias score included	21	0.41**	0.25, 0.57	5.03	39.86*	50
High risk of bias (>1)[a]	20	0.42**	0.26, 0.59	5.02	39.44*	52
Lowest risk of bias (0)[b]	15	0.45**	0.25, 0.65	4.47	39.98**	62
CBTp with individualized CF + primary outcome[c,d]						
Any risk of bias score included	16	0.51**	0.34, 0.68	5.99	23.15*	35
High risk of bias (>1)[a]	15	0.53**	0.36, 0.70	6.01	22.24	37
Lowest risk of bias (0)[b]	11	0.59**	0.39, 0.80	5.73	18.64*	46
Blinded RCTs only[e]						
All eligible CBTp RCTs	24	0.36**	0.20, 0.51	4.48	49.10**	53
Case formulation only	19	0.43**	0.26, 0.61	4.95	39.17**	54
Case formulation + primary outcome[c,d]	14	0.55**	0.37, 0.73	5.99	21.36	39
Additional analyses						
Group CBTp	4	0.11	−0.18, 0.41	0.76	3.15	5
Hallucinations as secondary outcome	5	0.05	−0.15, 0.24	0.46	3.87	0
Virtual-reality CBTp	2	0.56**	0.22, 0.89	3.27	0.75	0
Self-help CBTp	3	0.47	−0.42, 1.37	1.03	9.26**	78
After removal of 2 outliers						
Any risk of bias score included	26	0.27**	0.15, 0.40	4.37	34.35	27
High risk of bias (>1)[a]	24	0.27**	0.14, 0.39	4.16	32.52	29
Lowest risk of bias (0)[b]	16	0.31**	0.16, 0.46	4.05	24.21	39
Case formulation only	19	0.32**	0.19, 0.45	4.91	23.01	22
Case formulation + primary outcome[c,d]	14	0.41**	0.29, 0.54	6.56	9.64	0
Case formulation, primary outcome + RoB[a–c]	9	0.44**	0.31, 0.58	6.46	6.11	0
Excluding RCTs with high ratio nonschizophrenia spectrum						
Any risk of bias score included	26	0.32**	0.17, 0.47	4.23	49.46**	50
Lowest risk of bias (0)[b]	16	0.38**	0.19, 0.57	3.92	38.78**	61
Case formulation + primary outcome[c,d]	16	0.47**	0.30, 0.64	5.28	24.99	40
Case formulation, primary outcome + RoB[a–c]	10	0.58**	0.37, 0.79	5.35	17.58*	49

Note: All comparisons were using random model. Risk of bias scores refer to assessment using adapted version of the Cochrane Risk of Bias tool (0–4). CBTp, cognitive behavioral therapy for psychosis; CF, case formulation; CI, confidence interval; *g*, Hedges's *g*; n/a, not applicable; RCT, randomized controlled trial; RoB, risk of bias; TAU, treatment-as-usual.
Sensitivity analysis exclusions were as follows:
[a]Risk of bias score greater than 1 excluded: McLeod et al (2007); Peters et al (2010).
[b]Risk of bias score greater than 0 excluded: Cather et al (2005); Durham et al (2003); Gottlieb et al (2017); Krakvik et al (2013); Leff et al (2013); McLeod et al (2007). Peters et al (2010); Trouwer et al (2004); Wykes et al (2005).
[c]Hallucinations as primary outcome only: Birchwood et al (2014); Garety et al (2008); Haddock et al (2009); Tarrier et al (2014); Trouwer et al (2004).
[d]Case formulation only: Freeman et al (2015); Gottlieb et al (2017); Hazell et al (2017); McLeod et al (2007); Penn et al (2009); Wykes et al (2005).
[e]Nonblinded RCTs excluded: Krakvik et al (2013); McLeod et al (2007); Peters et al (2010).
*$P < .05$. **$P < .01$.

figures 3 and 4 provide a forest plot with all eligible RCTs included. When analyzing this broad sample of all 28 eligible RCTs for hallucinations, results demonstrated superiority of CBT over controls ($g = 0.34$, $P < .01$). When including only RCTs with the lowest possible risk of bias scores ($n = 19$), we observed a marginal but statistically nonsignificant increase in the magnitude of effect ($g = 0.40$, $P < .01$). Similarly, when including all RCTs comparing CBT against TAU, there was a significant effect favoring CBT ($g = 0.35$, $P < .01$), which had a marginal, nonsignificant increase when including only those RCTs with the lowest assessed risk ($n = 14$; $g = 0.41$, $P < .01$). The same pattern was observed when comparing CBT with active controls; CBT demonstrated superiority when all eligible RCTs were included ($g = 0.34$, $P < .01$), while including only the lowest risk RCTs resulted in a small and statistically nonsignificant increase in effect magnitude ($n = 5$; $g = 0.42$, $P < .01$).

We observed the same pattern when including only RCTs with hallucinations as the primary outcome target. When including all such RCTs, CBT demonstrated superiority over control ($g = 0.40$, $P < .02$) and increased when including only the lowest risk RCTs ($n = 14$; $g = 0.51$, $P < .02$). Similarly, when analyzing the impact of CBT with individually tailored case formulation vs controls we observed a significant effect when all eligible RCTs were included ($g = 0.41$, $P < .01$), and when including only the lowest risk RCTs ($n = 15$; $g = 0.45$, $p < .01$).

When including only blinded RCTs, CBT was superior to any control ($g = 0.36$, $P < .01$), when including only blinded case formulation RCTs ($n = 19$; $g = 0.43$, $P < .01$) and when limiting to RCTs, which applied blinded case formulation and hallucinations as primary outcome ($n = 14$; $g = 0.55$, $P < .01$).

When performing the most stringent comparison—namely including only RCTs, which utilized case formulation alongside targeting hallucinations as the primary outcome—we again observed the same pattern of increasing magnitude with bias reduction. The effect sizes demonstrated superiority for CBT when including all eligible RCTs ($g = 0.51$, $P < .01$) and the lowest bias risk RCTs ($n = 11$; $g = 0.59$, $P < .05$; see supplementary figure 5).

Effect of CBT on Delusions

Table 3 provides the results from all meta-analytic comparisons of CBT for delusions, whereas supplementary figure 4 provides a forest plot for all eligible RCTs. When including all eligible RCTs, CBT demonstrated superiority over controls ($g = 0.37$, $P < .01$). There was a marginal, nonsignificant reduction in the magnitude of this effect when including only the lowest risk RCTs ($g = 0.34$, $P < .01$). When including only comparisons against TAU, CBT demonstrated superiority against TAU when including all eligible RCTs ($g = 0.36$, $P < .01$), while a similar pattern of a small reduction of magnitude was present with the least risky RCTs ($g = 0.32$, $P < .01$). When comparing CBT with active controls, CBT did not demonstrate significant superiority when including all eligible RCTs ($g = 0.23$, $P = .16$) and when including only the lowest risk RCTs ($g = 0.30$, $P = .28$).

A similar pattern was present when including only RCTs with delusions as the primary outcome target; the magnitude of the significant effect in favor of CBT was highest when all eligible RCTs were included ($g = 0.38$, $P < .01$) and when including only the lowest risk RCTs ($g = 0.34$, $P < .01$). When including only RCTs with individually tailored case formulation, the effect size was consistent for the all eligible RCT comparison ($g = 0.37$, $P < .01$) and when including only the lowest risk RCTs ($g = 0.37$, $P < .01$).

When only blinded trials were included, CBT was superior to any control ($g = 0.31$, $P < .01$), which was consistent when limiting to case formulation RCTs ($g = 0.35$, $P < .01$) and RCTs with case formulation and delusions as primary outcome ($g = 0.34$, $P < .01$).

A similar pattern was observed in the most stringent comparison, which included only RCTs applying individualized case formulation with delusions as the primary outcome target. The effect favoring CBT was of highest magnitude when all eligible RCTs were included ($g = 0.38$, $P < .01$), whereas the effect was marginally lower when excluding RCTs with a high risk of bias ($g = 0.37$, $P < .01$) and RCTs with the lowest risk of bias ($g = 0.37$, $P < .01$).

Heterogeneity

There was a significant degree of heterogeneity present in the majority of comparisons. For hallucinations, the degree of significant heterogeneity ranged from 37% to 65%, indicating the existence of heterogeneity primarily within the moderate range across comparisons. Heterogeneity was lower in comparisons including RCTs, which utilized individualized case formulation and also targeted hallucinations as primary outcome focus. Heterogeneity in the delusions comparisons was overall higher, ranging from 39% to 75% and therefore indicating moderate to high heterogeneity. The sensitivity analyses for case formulation and primary outcome in the delusions category did not display a pattern of lower heterogeneity.

Publication Bias

The examination of funnel plots identified the possibility of unpublished negative studies across both symptom domains. For hallucinations, when all eligible RCTs were included, there was an estimation that 4 unpublished negative trials may exist. Duval and Tweedie's[54] trim and fill procedure provided an adjusted effect size by removing 5 RCTs. This procedure reduced the magnitude of the effect favoring CBT, but the effect remained significant ($g = 0.24$, 95% CI: 0.15–0.33). Egger's[55] test

Table 3. Effect Sizes of CBTp for Delusions

	N	*g*	95% CI	*Z*	*Q*-Value	I^2 (%)
Main comparison with all eligible RCTs						
Any risk of bias score included	27	0.37**	0.23, 0.52	4.95	54.54**	53
High risk of bias (>1)[a]	25	0.36**	0.20, 0.10	4.64	53.34**	55
Lowest risk of bias (0)[b]	18	0.34**	0.17, 0.50	4.02	39.33**	57
CBTp vs TAU						
Any risk of bias score included	22	0.36**	0.20, 0.52	4.34	48.96**	57
High risk of bias (>1)[a]	21	0.35**	0.18, 0.51	4.15	46.85**	57
Lowest risk of bias (0)[b]	16	0.32**	0.15, 0.49	3.64	34.42**	56
CBTp vs active intervention						
Any risk of bias score included	7	0.23	−0.19, 0.55	1.41	12.51	52
High risk of bias (>1)[a]	6	0.20	−0.45, 0.55	1.14	11.76	57
Lowest risk of bias (0)[b]	3	0.30	−0.25, 0.85	1.07	7.85*	75
CBTp with delusions as primary outcome[c]						
Any risk of bias score included	23	0.38**	0.22, 0.54	4.56	47.94**	54
High risk of bias (>1)[a]	21	0.36**	0.19, 0.53	4.21	45.83**	56
Lowest risk of bias (0)[b]	14	0.34**	0.15, 0.52	3.51	32.01**	59
CBTp with individualized case formulation[d]						
Any risk of bias score included	21	0.37**	0.20, 0.54	4.33	49.43**	60
High risk of bias (>1)[a]	20	0.37**	0.20, 0.54	4.19	49.10**	61
Lowest risk of bias (0)[b]	16	0.37**	0.19, 0.55	4.00	38.09**	61
CBTp with individualized CF + primary outcome[c,d]						
Any risk of bias score included	17	0.38**	0.19, 0.57	3.86	41.93**	62
High risk of bias (>1)[a]	16	0.37**	0.18, 0.57	3.70	41.60**	64
Lowest risk of bias (0)[b]	12	0.37**	0.67, 0.58	3.48	30.68**	64
Blinded RCTs only[e]						
All eligible CBTp RCTs	22	0.31**	0.16, 0.47	3.96	46.77**	55
Case formulation only	19	0.35**	0.17, 0.52	3.90	45.44**	60
Case formulation + primary outcome[c,d]	15	0.34**	0.14, 0.54	3.37	38.11**	63
Additional analyses						
Group CBTp	2	0.35	−0.02, 0.72	1.84	0.74	0
Delusions as secondary outcome	4	0.36	−0.06, 0.78	1.69	7,15	58
Virtual-reality CBTp	2	0.56**	0.24, 0.89	3.36	0.86	0
After removal of 1 outlier						
Any risk of bias score included	26	0.32**	0.19, 0.46	4.71	41.30*	39
High risk of bias (>1)[a]	24	0.31**	0.17, 0.44	4.41	38.72*	41
Lowest risk of bias (0)[b]	17	0.26**	0.13, 0.40	3.81	23.96	33
Case formulation only	20	0.31**	0.16, 0.47	4.07	34.90*	46
Case formulation + primary outcome[c,d]	16	0.31**	0.14, 0.498	3.59	27.72*	46
Case formulation, primary outcome + RoB[a–c]	11	0.28**	0.11, 0.45	3.26	15.99	37

Note: All comparisons were using random model. Risk of bias scores refer to assessment using adapted version of the Cochrane Risk of Bias tool (0–4). CBTp, cognitive behavioral therapy for psychosis; CF, case formulation; CI, confidence interval; *g*, Hedges's *g*; n/a, not applicable; RCT, randomized controlled trial; RoB, risk of bias; TAU, treatment-as-usual.
Sensitivity analysis exclusions were as follows:
[a]Risk of bias score greater than 1 excluded: Foster et al (2010); O'Connor et al (2007).
[b]Risk of bias score greater than 0 excluded: Cather et al (2005); Durham et al (2003); Freeman et al (2016); Gottlieb et al (2017); Krakvik et al (2013); Foster et al (2010); O'Connor et al (2017); Waller et al (2015).
[c]Hallucinations as primary outcome only: Freeman et al (2014); Garety et al (2008); Haddock et al (2009); Tarrier et al (2014).
[d]Case formulation only: Foster et al (2010); Freeman et al (2015); Gottlieb et al (2017); Penn et al (2009); Waller et al (2015).
[e]Nonblinded RCTs excluded: Foster et al (2010); Freeman et al (2016); Krakvik et al (2013); O'Connor et al (2007); Waller et al (2015).
*$P < .05$. **$P < .01$.

of the intercept was not significant, whereas the classic fail-safe N estimate that 291 unpublished studies would have to exist to bring the *P*-value above the alpha level of .05. For delusions, the funnel plot estimated the existence of 8 unpublished trials. The trim and fill procedure provided an adjusted effect removing 7 RCTs, which again remained significant although had reduced magnitude ($g = 0.18$, 95% CI: 0.09–0.26). Egger's test of the intercept was significant on this comparison, whereas the classic fail-safe N suggested that it would require 335 missing studies to bring the *P*-value to above the .05 alpha level.

Post Hoc Investigation of Outliers

We identified significant heterogeneity across a high proportion of comparisons for auditory hallucinations

that was not observed in the previous review. We therefore examined forest plots to identify primary studies as potential outliers contributing to high heterogeneity. Examination of supplementary figure 3 suggested that the trial by Habib et al[47] was a significant outlier because its 95% CI did not overlap with that of the pooled effect size. The effect from one RCT by Naeem et al[49] was also identified as a potential outlier. We therefore assessed heterogeneity when excluding both outliers in an exploratory sensitivity analysis. Excluding both RCTs reduced the heterogeneity in the comparison including all eligible RCTs below the alpha .05 level to $I = 27\%$ ($Q = 33.35$, $P = .10$). Heterogeneity was gradually reduced in subsequent sensitivity analyses and was observed as 0% in the most stringent and homogenous group of RCTs (case formulation, hallucinations as primary outcome, and minimal bias risk). We also investigated the possible impact of the outliers on the magnitude of effects. We observed nonsignificant reduction in the effect magnitude across categories although the pattern of marginally increasing magnitude following stricter sensitivity analyses was maintained. Results from outlier exclusion for hallucinations are reported in table 2.

Similar examination of CIs in supplementary figure 4 identified the effect size from the Naeem et al[46] trial in delusions as an outlier. We therefore completed the same set of sensitivity analyses when excluding this RCT. Results demonstrated that heterogeneity was broadly reduced; in some comparisons to the extent that heterogeneity was no longer significant. There were also marginal and statistically insignificant reductions in the effect size. Results from outlier exclusion for delusions are reported in table 3.

Post Hoc Sensitivity Analyses

The length of treatment varied considerably between RCTs; the shortest treatment was a single session of VR-CBTp (Freeman et al, 2016), whereas the longest CBTp treatments lasted 9 months. To investigate the impact of this variation, we completed 2 further post hoc analyses: first, a sensitivity analysis excluding the shortest treatment in the delusions comparison and second a meta-regression investigating the impact of treatment length on effects for both hallucinations and delusions. The results of the sensitivity analyses are reported in tables 2 and 3. Removal of the shortest RCT resulted in only marginal changes to effect sizes. The meta-analysis showed that the number of CBTp session participants received did not have a significant impact on the effect for hallucinations ($P = .88$) or delusions ($P = .63$). These findings were consistent when controlling for risk of bias.

We also conducted a sensitivity analyses when removing 2 RCTs with a higher proportion of participants with psychosis diagnosed with nonschizophrenia-spectrum disorders.[38,39] Results of these sensitivity analyses are presented in table 2 and show only marginal changes to effect sizes.

Post Hoc Case Formulation Head-to-Head Comparison

We completed a direct comparison of the RCTs including CBTp case formulation vs those without. In the delusions analysis, CBTp demonstrated significant effects of similar magnitude for case formulation trials ($g = 0.38$, $P < .01$) and noncase formulation trials ($g = 0.35$, $P < .05$). In the hallucinations analysis, CBTp demonstrated a significant effect for case formulation trials ($g = 0.40$, $P < .5$), but not for noncase formulation trials ($g = 0.10$, $P = .51$).

Cumulative Meta-analysis

Figure 2 depicts the cumulative forest plots for both the CBTp for hallucinations and delusions comparisons when including all eligible RCTs. This figure demonstrates the stability of the effect size over time. Table 4 (Supplementary materials) provides fail-safe ratio calculations for all 4 cumulative meta-analyses; namely the main analysis comparisons for both hallucinations and delusions when including all eligible RCTs alongside the most stringent sensitivity analysis when including only RCTs that scored zero on the risk of bias assessment, utilized individualized case formulation and had primary outcome focus. More extensive figures for all cumulative meta-analyses including all relevant data are available in supplementary figures 6–9.

For hallucinations, the 1.0 level of the fail-safe ratio demonstrating sufficiency was surpassed in 2016, which was consistent in the sensitivity analysis. For delusions, the 1.0 level was surpassed in 2015 for the main analysis and in 2017 for the sensitivity analysis. Cumulative forest plots for each of the remaining 3 comparisons are included in supplementary materials and demonstrate stability of the effect size.

Discussion

Cumulative Meta-analysis: Sufficient and Stable

This cumulative meta-analysis allowed us to demonstrate that the existing evidence base for the effect of CBTp on hallucinations and delusions is both statistically stable and sufficient according to Muellerleile and Mullen's[11] guidelines. A notable demonstration of the stability of the evidence base is that the addition of a large trial with a null finding[53] had only a marginal impact on the effect size ($g = 0.358$ to $g = 0.351$ for hallucinations and $g = 0.383$ to $g = 0.363$ for delusions). The evidence base for hallucinations has been sufficient since 2016, after which another 6 RCTs were added. Similarly, our review suggests sufficiency of evidence for delusions from 2015 after which point 6 RCTs have also contributed data. Our findings suggest that further RCTs repeatedly testing

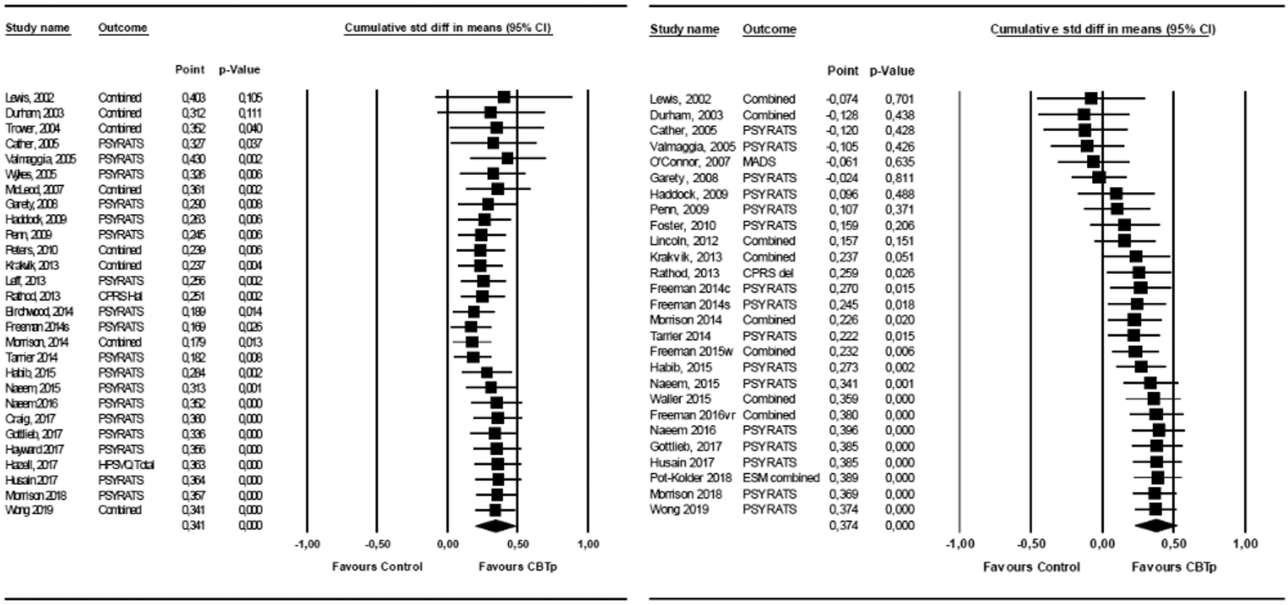

Fig. 2. Cumulative meta-analysis forest plots for CBTp for (a) hallucinations and (b) delusions, all eligible RCTs.

CBTp are unlikely to have a significant impact on the magnitude or significance of treatment effects or to alter our conclusions in any substantive way, although we note that in conventional meta-analysis CBTp did not demonstrate superiority for delusions compared with active controls in the context of low power.

Conventional Meta-analysis

The conventional meta-analytic comparisons in this review provided broadly similar results to our earlier review[1] despite adding 19 RCTs published during the 6 years elapsed since the previous systematic search. There were however notable differences in some comparisons. For hallucinations, when including only RCTs utilizing both case formulation and primary outcome focus, the effect size increased to $g = 0.6$ when controlling for risk of bias. However, after removing 2 outliers, this effect shrank to $g = 0.44$, which is consistent with our 2014 review. We observed a broadly consistent pattern across comparisons for hallucinations; when risk of bias was minimized and when including only case formulation and primary outcome focus, the magnitude of effects increased marginally but not significantly. Effects remained in the range of $g = 0.3$ to $g = 0.6$. The facility to examine risk of bias in this specific form of sensitivity analysis was not included in the previous review; therefore, this finding, alongside the broad consistency of results in the hallucinations domain, further suggests robust evidence of the impact of targeted, formulation-driven CBTp for hallucinations.

The effects of CBTp on delusions were of similar magnitude to those for auditory hallucinations when including all eligible trials, although did not display the pattern of marginally increasing magnitude when excluding RCTs with a higher risk of bias. Effect sizes in delusions comparisons remained in the region of $g = 0.32$–0.38 for all main comparisons with the exception of the nonsignificant comparisons against active treatments. It should be noted that this category was comparatively underpowered and that the sensitivity analysis that included only RCTs with the lowest bias risk provided a significant effect of a similar magnitude ($g = 0.3$, $P < .05$). Despite the finding that CBTp was not superior to active control treatments for delusions, because CBTp for delusions was demonstrated as meta-analytically effective overall, while the active control conditions have no meta-analytical evidence, we suggest that CBTp for delusions continues to be recommended until evidence for other treatments emerges.

Our head-to-head comparison of case formulation–driven CBTp compared with that without also suggests that case formulation–driven CBTp is more effective in reducing hallucinations, whereas no difference was evident in the effects for delusions. We note that there were significantly more RCTs in the case formulation arm and therefore lower power in the noncase formulation arm, although the lower effect magnitude for nonformulation-based CBTp for hallucinations is still indicative of potential inferiority. We note a recent secondary analysis[57] of one RCT included in our review,[32] which failed to find a significant effect of case formulation on outcome. This study also reported a nonsignificant trend of poorer treatment outcome for case formulation participants. Our findings are on a meta-analytic level indicative that case formulation is more beneficial for hallucinations, although definitive comment awaits more RCTs becoming available in the noncase formulation arm. Because many novel CBTp applications adhere less to the traditional formulation-based treatment approach, further pooling and comparison of this developing dichotomy is warranted.

Limitations

A notable limitation in this meta-analytic review was significant heterogeneity across a high proportion of the comparisons. Significant heterogeneity was present only in comparisons for delusions in the previous 2014 review; no hallucinations comparison in the original review demonstrated significant heterogeneity. Post hoc investigation established that heterogeneity introduced to the hallucinations comparisons was largely attributable to 2 outliers, one of which adapted CBTp for application in other cultural settings,[47] while the other applied group-based self-help CBTp.[49] Similarly, another RCT of culturally adapted CBTp contributed to heterogeneity in the delusions comparisons.[58] Our earlier review conceptualized case formulation–driven CBTp RCTs as "apples" in comparison to "oranges"; a broader and more inclusive sample of RCTs applying CBTp principles in alternative style. We may therefore consider the newer, less homogenous CBTp trials and interventions again as such "oranges." The development of such novel approaches and application across wider settings is of importance in the CBTp field; therefore, we expect further such heterogeneity in future reviews. We also acknowledge that a number of comparisons in our review—namely those examining novel interventions and those comparing CBTp to active interventions—were underpowered. Low power therefore means that there exists potential for Type 2 error in missing effects that do exist. We also acknowledge the limitation of our narrow focus relying only on pre–post change, meaning that we cannot report on enduring effects at longer-term follow-up. Our focus on the specific hallucination and delusion outcomes also meant that other important outcomes such as relapse, functioning, or level of distress were not considered, while focusing on schizophrenia-spectrum diagnoses also excludes many experiencing psychosis as a symptom of other diagnoses such as bipolar disorder and substance use disorders.[59,60]

Future Research

Our cumulative meta-analysis suggests that there is little value in researchers repeatedly testing conventional, formulation-driven CBTp in further RCTs; because the evidence base has demonstrated sufficiency, resources may better be directed toward novel approaches. The question of whether CBTp "works" is no longer central, while previous disputes appear to have been settled.[6] Further development of RCTs examining novel approaches such as culturally adapted CBTp and VR-CBTp will allow clearer conclusion on their efficacy via increased power in meta-analysis including only these interventions. We also note that RCTs examining novel approaches typically provide briefer interventions, although our post hoc analysis did not suggest a significant impact of treatment duration on outcome. Our results also suggest that there may be limited value in "collecting" further conventional meta-analyses, which Murray[58] notably compared with the hobbyist pursuit of postage stamps. There is however the possibility that individual-participant data meta-analysis techniques may be applied by combining the original databases of CBTp RCTs to provide more precise estimation of effects and the examination of moderating variables (eg, demographic or clinical characteristics) on specific hallucination and delusion outcomes. Due to the identification of potential publication bias, we also encourage any researchers contributing to the "file drawer problem" to publish any relevant trials, which are not yet available in the public sphere for meta-analytic comparison. Future research may also focus further on the intricacies of the relationship between CBTp and antipsychotics; despite the demonstrated sufficiency of evidence for CBTp, it remains to date investigated primarily as an adjunctive treatment.[61] Finally, although interesting findings such as the pattern of increasing effect magnitude when primary outcome focus or case formulation are applied, definitive comment on the effectiveness of specific CBTp components awaits detailed dismantling studies. There may therefore be opportunity to apply the developing factorial design principles of intervention optimization research to the psychosis field.[62]

Conclusions

This meta-analytic review further demonstrates the efficacy of CBTp for auditory hallucinations and delusions and suggests that the evidence base is now both sufficient and stable. The robust performance of the effect on hallucinations in sensitivity analyses supports the notion that CBTp is particularly effective in this domain, whereas heterogeneity and potential publication bias are issues that should be carefully examined in future reviews as further research becomes available for inclusion.

Supplementary Material

Supplementary data are found at *Schizophrenia Bulletin* online.

Funding

None.

Acknowledgments

We acknowledge Professor Pim Cuijpers for supporting this project. The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References

1. van der Gaag M, Valmaggia LR, Smit F. The effects of individually tailored formulation-based cognitive behavioural therapy in auditory hallucinations and delusions: a meta-analysis. *Schizophr Res.* 2014;156(1):30–37.

2. Turner DT, van der Gaag M, Karyotaki E, Cuijpers P. Psychological interventions for psychosis: a meta-analysis of comparative outcome studies. *Am J Psychiatry.* 2014;171(5):523–538.
3. Jauhar S, McKenna PJ, Radua J, Fung E, Salvador R, Laws KR. Cognitive-behavioural therapy for the symptoms of schizophrenia: systematic review and meta-analysis with examination of potential bias. *Br J Psychiatry.* 2014;204(1):20–29.
4. Husain MO, Chaudhry IB, Mehmood N, et al. Pilot randomised controlled trial of culturally adapted cognitive behavior therapy for psychosis (CaCBTp) in Pakistan. *BMC Health Serv Res.* 2017;17(1):808.
5. Pot-Kolder RMCA, Geraets CNW, Veling W, et al. Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: a single-blind randomised controlled trial. *Lancet Psychiatry.* 2018;5(3):217–226.
6. McKenna P, Leucht S, Jauhar S, Laws K, Bighelli I. The controversy about cognitive behavioural therapy for schizophrenia. *World Psychiatry.* 2019;18(2):235–236.
7. Kuipers E, Yesufu-Udechuku A, Taylor C, Kendall T. Management of psychosis and schizophrenia in adults: summary of updated NICE guidance. *BMJ.* 2014;348:g1173.
8. Cooke A. *Understanding Psychosis and Schizophrenia: Why People Sometimes Hear Voices, Believe Things That Others Find Strange, or Appear out of Touch With Reality… and What Can Help*. London, UK: British Psychological Society; 2017.
9. Jones C, Hacker D, Xia J, et al. Cognitive behavioural therapy plus standard care versus standard care for people with schizophrenia. *Cochrane Database Syst Rev.* 2018;2018(12):1–227. doi:10.1002/14651858.CD007964.pub2
10. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med.* 1992;327(4):248–254.
11. Muellerleile P, Mullen B. Sufficiency and stability of evidence for public health interventions using cumulative meta-analysis. *Am J Public Health.* 2006;96(3):515–522.
12. Love R, Adams J, van Sluijs EMF, Foster C, Humphreys D. A cumulative meta-analysis of the effects of individual physical activity interventions targeting healthy adults. *Obes Rev.* 2018;19(8):1164–1172.
13. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1–e34.
14. Chadwick P, Strauss C, Jones AM, et al. Group mindfulness-based intervention for distressing voices: a pragmatic randomised controlled trial. *Schizophr Res.* 2016;175(1–3):168–173.
15. Zimmermann G, Favrod J, Trieu VH, Pomini V. The effect of cognitive behavioral treatment on the positive symptoms of schizophrenia spectrum disorders: a meta-analysis. *Schizophr Res.* 2005;77(1):1–9.
16. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr Bull.* 1987;13(2):261–276.
17. Haddock G, Barrowclough C, Shaw JJ, Dunn G, Novaco RW, Tarrier N. Cognitive-behavioural therapy v. social activity therapy for people with psychosis and a history of violence: randomised controlled trial. *Br J Psychiatry.* 2009;194(2):152–157.
18. Steel C, Garety PA, Freeman D, et al. The multidimensional measurement of the positive symptoms of psychosis. *Int J Methods Psychiatr Res.* 2007;16(2):88–96.
19. Higgins JPT. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557–560.
20. Turner DT, McGlanaghy E, Cuijpers P, van der Gaag M, Karyotaki E, MacBeth A. A meta-analysis of social skills training and related interventions for psychosis. *Schizophr Bull.* 2018;44(3):475–491.
21. Jackson D, Turner R. Power analysis for random-effects meta-analysis. *Res Synth Methods.* 2017;8(3):290–302.
22. Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull.* 1979;86(3):638–641.
23. Cuijpers P. *Meta-analyses in Mental Health Research – A Practical Guide.* Vol. 15; Amsterdam, Netherlands: Pim Cuijpers Uitgeverij; 2016. doi:10.1016/j.clinthera.2009.11.030.
24. Durham RC, Guthrie A, Morton RV, et al. Tayside-Fife clinical trial of cognitive-behavioural therapy formedication-resistant psychotic symptoms: results to 3-month follow-up. Br J Psychiatry. 1992;182:303–312.
25. Valmaggia LR, van der Gaag M, Tarrier N, Pijnenborg M, Slooff CJ. Cognitive-behavioural therapy for refractory psychotic symptoms of schizophrenia resistant to atypical antipsychotic medication. Randomised controlled trial. *Br J Psychiatry.* 2005;186:324–330.
26. Connor KO, Stip E, Pelissier M, et al. Treating delusional disorder: a comparison of attention placebo control. *Revue.* 2007;52(3):182–190.
27. Penn DL, Meyer PS, Evans E, Wirth RJ, Cai K, Burchinal M. A randomized controlled trial of group cognitive-behavioral therapy vs. enhanced supportive therapy for auditory hallucinations. *Schizophr Res.* 2009;109(1–3):52–59.
28. Craig TK, Rus-Calafell M, Ward T, et al. AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial. *Lancet Psychiatry.* 2018;5(1):31–40.
29. Lewis S, Tarrier N, Haddock G, et al. Randomised controlled trial of cognitive-behavioural therapy in early schizophrenia: acute-phase outcomes. *Br J Psychiatry.* 2002;181(suppl. 43):s91–s97.
30. Cather C, Penn D, Otto MW, Yovel I, Mueser KT, Goff DC. A pilot study of functional cognitive behavioral therapy (fCBT) for schizophrenia. *Schizophr Res.* 2005;74(2–3):201–209.
31. Freeman D, Bradley J, Antley A, et al. Virtual reality in the treatment of persecutory delusions: randomised controlled experimental study testing how to reduce delusional conviction. *Br J Psychiatry.* 2016;209(1):62–67.
32. Morrison AP, Turkington D, Pyle M, et al. Cognitive therapy for people with schizophrenia spectrum disorders not taking antipsychotic medication: a randomised controlled trial. *Schizophr Res.* 2014;153:S75.
33. Trower P, Birchwood M, Meaden A, Byrne S, Nelson A, Ross K. Cognitive therapy for command hallucinations: randomised controlled trial. *Br J Psychiatry.* 2004;184:312–320.
34. Surguladze S, Fannon D, Wykes T, et al. What are the effects of group cognitive behaviour therapy for voices? A randomised control trial. *Schizophr Res.* 2005;77(2–3):201–210.
35. Mcleod T, Morris M, Birchwood M, Dovey A. Work with voice hearers. Part 1. *Br J Nurs.* 2007;16(4):248–252.
36. Garety PA, Fowler DG, Freeman D, Bebbington P, Dunn G, Kuipers E. Cognitive–behavioural therapy and family intervention for relapse prevention and symptom reduction in

psychosis: randomised controlled trial. *Br J Psychiatry.* 2008;192(6):412–423.

37. Foster C, Startup H, Potts L, Freeman D. A randomised controlled trial of a worry intervention for individuals with persistent persecutory delusions. *J Behav Ther Exp Psychiatry.* 2010;41(1):45–51.
38. Lincoln TM, Ziegler M, Mehl S, et al. Moving from efficacy to effectiveness in cognitive behavioral therapy for psychosis: a randomized clinical practice trial. *J Consult Clin Psychol.* 2012;80(4):674–686.
39. Kråkvik B, Gråwe RW, Hagen R, Stiles TC. Cognitive behaviour therapy for psychotic symptoms: a randomized controlled effectiveness trial. *Behav Cogn Psychother.* 2013;41(5):511–524.
40. Rathod S, Phiri P, Harris S, et al. Cognitive behaviour therapy for psychosis can be adapted for minority ethnic groups: a randomised controlled trial. *Schizophr Res.* 2013;143(2–3):319–326.
41. Leff J, Williams G, Huckvale MA, Arbuthnot M, Leff AP. Computer-assisted therapy for medication-resistant auditory hallucinations: proof-of-concept study. *Br J Psychiatry.* 2013;202:428–433.
42. Birchwood M, Michail M, Meaden A, et al. Cognitive behaviour therapy to prevent harmful compliance with command hallucinations (COMMAND): a randomised controlled trial. *Lancet Psychiatry.* 2014;1(1):23–33.
43. Freeman D, Pugh K, Dunn G, et al. An early phase II randomised controlled trial testing the effect on persecutory delusions of using CBT to reduce negative cognitions about the self: the potential benefits of enhancing self confidence. *Schizophr Res.* 2014;160(1–3):186–192.
44. Tarrier N, Kelly J, Maqsood S, et al. The cognitive behavioural prevention of suicide in psychosis: a clinical trial. *Schizophr Res.* 2014;156(2–3):204–210.
45. Freeman D, Waite F, Startup H, et al. Efficacy of cognitive behavioural therapy for sleep improvement in patients with persistent delusions and hallucinations (BEST): a prospective, assessor-blind, randomised controlled pilot trial. *Lancet Psychiatry.* 2015;2(11):975–983.
46. Naeem F, Saeed S, Irfan M, et al. Brief culturally adapted CBT for psychosis (CaCBTp): a randomized controlled trial from a low income country. *Schizophr Res.* 2015;164(1–3):143–148.
47. Habib N, Dawood S, Kingdon D, Naeem F. Preliminary evaluation of culturally adapted CBT for psychosis (CA-CBTp): findings from developing culturally-sensitive CBT project (DCCP). *Behav Cogn Psychother.* 2015;43(2):200–208.
48. Waller H, Emsley R, Freeman D, et al. Thinking well: a randomised controlled feasibility study of a new CBT therapy targeting reasoning biases in people with distressing persecutory delusional beliefs. *J Behav Ther Exp Psychiatry.* 2015;48:82–89.
49. Naeem F, Johal R, McKenna C, et al. Cognitive behavior therapy for psychosis based guided self-help (CBTp-GSH) delivered by frontline mental health professionals: results of a feasibility study. *Schizophr Res.* 2016;173(1–2):69–74.
50. Hayward M, Jones AM, Bogen-Johnston L, Thomas N, Strauss C. Relating therapy for distressing auditory hallucinations: a pilot randomized controlled trial. *Schizophr Res.* 2017;183:137–142.
51. Hazell CM, Hayward M, Cavanagh K, Jones AM, Strauss C. Guided self-help cognitive-behaviour Intervention for VoicEs (GiVE): results from a pilot randomised controlled trial in a transdiagnostic sample. *Schizophr Res.* 2018;195:441–447.
52. Gottlieb JD, Gidugu V, Maru M, et al. Randomized controlled trial of an internet cognitive behavioral skills-based program for auditory hallucinations in persons with psychosis. *Psychiatr Rehabil J.* 2017;40(3):283–292.
53. Morrison AP, Pyle M, Gumley A, et al.; FOCUS Trial Group. Cognitive behavioural therapy in clozapine-resistant schizophrenia (FOCUS): an assessor-blinded, randomised controlled trial. *Lancet Psychiatry.* 2018;5(8):633–643.
54. Wong AWS, Ting KT, Chen EYH. Group cognitive behavioural therapy for Chinese patients with psychotic disorder: a feasibility controlled study. *Asian J Psychiatr.* 2019;39:157–164.
55. 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(2):455–463.
56. Egger M, Smith GD. Bias in meta-analysis detected by a simple, graphical test measures of funnel plot asymmetry. *BMJ.* 2011;315(7109):1–21.
57. Spencer HM, McMenamin M, Emsley R, et al. Cognitive behavioral therapy for antipsychotic-free schizophrenia spectrum disorders: does therapy dose influence outcome? *Schizophr Res.* 2018;202:385–386.
58. Murray RM. On collecting meta-analyses of schizophrenia and postage stamps. *Psychol Med.* 2014;44(16):3407–3408.
59. Keck PE Jr, McElroy SL, Havens JR, et al. Psychosis in bipolar disorder: phenomenology and impact on morbidity and course of illness. *Compr Psychiatry.* 2003;44(4):263–269.
60. Voce A, Calabria B, Burns R, Castle D, McKetin R. A systematic review of the symptom profile and course of methamphetamine-associated psychosis substance use and misuse. *Subst Use Misuse.* 2019;54(4):549–559.
61. Bighelli I, Salanti G, Huhn M, et al. Psychological interventions to reduce positive symptoms in schizophrenia: systematic review and network meta-analysis. *World Psychiatry.* 2018;17(3):316–329.
62. Collins L. *Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimisation Strategy (MOST)*. New York, NY: Springer; 2018.