JAMA Psychiatry | Original Investigation

Long-term Outcomes of Cognitive Behavioral Therapy for Anxiety-Related Disorders
A Systematic Review and Meta-analysis

Eva A. M. van Dis, MSc; Suzanne C. van Veen, MSc; Muriel A. Hagenaars, PhD; Neeltje M. Batelaan, PhD; Claudi L. H. Bockting, PhD; Rinske M. van den Heuvel, MSc; Pim Cuijpers, PhD; Iris M. Engelhard, PhD

IMPORTANCE Cognitive behavioral therapy is recommended for anxiety-related disorders, but evidence for its long-term outcome is limited.

OBJECTIVE This systematic review and meta-analysis aimed to assess the long-term outcomes after cognitive behavioral therapy (compared with care as usual, relaxation, psychoeducation, pill placebo, supportive therapy, or waiting list) for anxiety disorders, posttraumatic stress disorder (PTSD), and obsessive-compulsive disorder (OCD).

DATA SOURCES English-language publications were identified from PubMed, PsycINFO, Embase, Cochrane, OpenGrey (1980 to January 2019), and recent reviews. The search strategy included a combination of terms associated with anxiety disorders (eg, *panic* or *phobi**) and study design (eg, *clinical trial* or *randomized controlled trial*).

STUDY SELECTION Randomized clinical trials on posttreatment and at least 1-month follow-up effects of cognitive behavioral therapy compared with control conditions among adults with generalized anxiety disorder, panic disorder with or without agoraphobia, social anxiety disorder, specific phobia, PTSD, or OCD.

DATA EXTRACTION AND SYNTHESIS Researchers independently screened records, extracted statistics, and assessed study quality. Data were pooled using a random-effects model.

MAIN OUTCOMES AND MEASURES Hedges *g* was calculated for anxiety symptoms immediately after treatment and at 1 to 6 months, 6 to 12 months, and 12 months or more after treatment completion.

RESULTS Of 69 randomized clinical trials (4118 outpatients) that were mainly of low quality, cognitive behavioral therapy compared with control conditions was associated with improved outcomes after treatment completion and at 1 to 6 months and at 6 to 12 months of follow-up for a generalized anxiety disorder (Hedges *g*, 0.07-0.40), panic disorder with or without agoraphobia (Hedges *g*, 0.22-0.35), social anxiety disorder (Hedges *g*, 0.34-0.60), specific phobia (Hedges *g*, 0.49-0.72), PTSD (Hedges *g*, 0.59-0.72), and OCD (Hedges *g*, 0.70-0.85). At a follow-up of 12 months or more, these associations were still significant for generalized anxiety disorder (Hedges *g*, 0.22; number of studies [*k*] = 10), social anxiety disorder (Hedges *g*, 0.42; *k* = 3), and PTSD (Hedges *g*, 0.84; *k* = 5), but not for panic disorder with or without agoraphobia (*k* = 5) and could not be calculated for specific phobia (*k* = 1) and OCD (*k* = 0). Relapse rates after 3 to 12 months were 0% to 14% but were reported in only 6 randomized clinical trials (predominantly for panic disorder with or without agoraphobia).

CONCLUSIONS AND RELEVANCE The findings of this meta-analysis suggest that cognitive behavioral therapy for anxiety-related disorders is associated with improved outcomes compared with control conditions until 12 months after treatment completion. At a follow-up of 12 months or more, effects were small to medium for generalized anxiety disorder and social anxiety disorder, large for PTSD, and not significant or not available for other disorders. High-quality randomized clinical trials with 12 months or more of follow-up and reported relapse rates are needed.

JAMA Psychiatry. 2020;77(3):265-273. doi:10.1001/jamapsychiatry.2019.3986
Published online November 23, 2019. Corrected on May 13, 2020.

Supplemental content

Author Affiliations: Department of Clinical Psychology, Utrecht University, Utrecht, the Netherlands (van Dis, van Veen, Hagenaars, van den Heuvel, Engelhard); Department of Psychiatry, Amsterdam Universitair Medisch Centrum, location Vrije Universiteit medisch centrum, Amsterdam, the Netherlands (Batelaan); Department of Psychiatry, Amsterdam Universitair Medisch Centrum, location Academisch Medisch Centrum, University of Amsterdam, Amsterdam, the Netherlands (Bockting); Department of Clinical, Neuro, and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands (Cuijpers).

Corresponding Author: Eva A. M. van Dis, MSc, Department of Clinical Psychology, Utrecht University, PO Box 80140, 3508 TC Utrecht, the Netherlands (e.a.m.vandis@uu.nl).

Anxiety disorders, posttraumatic stress disorder (PTSD), and obsessive-compulsive disorder (OCD) are highly prevalent[1,2] and are associated with substantial personal[3] and societal costs.[4-6] Clinical practice guidelines recommend psychological and pharmacological interventions for anxiety-related disorders,[7-13] but most patients favor psychotherapy over pharmacotherapy.[14] Cognitive behavioral therapy (CBT) for these disorders has been associated with reduced symptoms at short term,[15,16] with small to medium effect sizes adjusted for publication bias and when studies with waiting list comparisons were not taken into account.[15] However, regarding its long-term outcome, little meta-analytic evidence is available. Such evidence is important because the course of anxiety-related disorders is typically chronic.[17] Evidence on the long-term outcome is particularly vital for researchers to prioritize research directions (eg, further examining variables associated with treatment success and ways to optimize treatment) and for clinicians to give patients realistic information.

Four recent meta-analyses have addressed the long-term outcome of CBT for anxiety-related disorders, and they generally indicate a medium symptom reduction up to 2 years following treatment completion.[18-21] However, in 2 of these,[18,21] CBT outcome was only calculated over time (pretreatment vs posttreatment vs follow-up) and not relative to a control condition. Therefore, these meta-analyses could not disentangle treatment outcome from placebo effects or spontaneous remission. Moreover, because pretreatment and posttreatment correlations of individual studies are often unknown, there may be substantial errors in these effect size estimations.[22] The other 2 meta-analyses did use control conditions, but these were limited to placebo,[19] resulting in 23 studies, or relaxation,[20] resulting in 27 studies. The numbers of studies would be at least twice as large if other comparison conditions were also included (eg, a care-as-usual group). In addition, no meta-analysis has examined the association between CBT and relapse rates in anxiety-related disorders, to our knowledge. Cross-sectional findings indicate that approximately 31% to 55% of patients with remitted anxiety meet diagnostic criteria of the same or another disorder within 4 years.[23] Research on relapse and the return of fear has become a major focus of fundamental fear and anxiety research,[24] but the evidence for clinical relapse after psychotherapy in anxiety-related disorders is limited.

Our aim was to conduct a comprehensive meta-analysis to establish a reliable estimate of the long-term outcome of CBT relative to passive and active comparison groups in anxiety disorders, PTSD, and OCD. We examined (1) long-term effects (≥1-month posttreatment) and (2) relapse rates after successful treatment in patients with generalized anxiety disorder (GAD), panic disorder with or without agoraphobia (PD), social anxiety disorder (SAD), specific phobia, PTSD, and OCD.

Key Points

Question What is the long-term outcome of cognitive behavioral therapy for anxiety disorders, posttraumatic stress disorder, and obsessive-compulsive disorder?

Findings In this systematic review and meta-analysis of 69 randomized clinical trials including 4118 patients, cognitive behavioral therapy was associated with better outcomes compared with control conditions among patients with anxiety symptoms within 12 months after treatment completion. At longer follow-up, significant associations were found only for generalized anxiety disorder, social anxiety disorder, and posttraumatic stress disorder; relapse rates (predominantly for panic disorder with or without agoraphobia) after 3 to 12 months were 0% to 14%.

Meaning The findings suggest that compared with control conditions, cognitive behavioral therapy was generally associated with lower anxiety symptoms within 12 months after treatment completion, but few studies have examined longer-term outcomes.

Methods

The systematic review and meta-analysis was preregistered at PROSPERO,[25] and it adhered to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline.[26]

Search Strategy

Relevant English-language publications were identified by systematically searching PubMed, PsycINFO, Embase, Cochrane, and OpenGrey (from 1980 until January 2019). The search strategy included a combination of terms related to anxiety disorders (eg, *panic* or *phobi**) and study design (eg, *clinical trial* or *randomized controlled trial*). eTable 1 in the Supplement provides the exact search strategies. The electronic database search was supplemented with a bibliography screening of 4 relevant meta-analyses[18-21] and 1 systematic review.[27]

Inclusion Criteria

Randomized clinical trials were included that examined effects of CBT (ie, any therapy with cognitive restructuring and/or a behavioral therapy, such as exposure, as core component),[15] including third generation CBTs (ie, acceptance and commitment therapy and metacognitive therapy), at least 1 month after treatment completion, in an individual, group, or internet treatment format. Comparison groups included care as usual (ie, anything patients would normally receive as long as it was not a structured type of psychotherapy, such as primary care at medical centers or case management with educational groups),[15] relaxation, psychoeducation, pill placebo, supportive therapy, or waiting list. Studies were included if they tested adult patients (or samples consisting mostly of adults but also some adolescents aged ≥16 years) who received a diagnosis of GAD, PD, SAD, specific phobia, PTSD, or OCD based on results of a structured diagnostic interview.

Studies were excluded if they did not use CBT (eg, applied relaxation, eye movement desensitization and reprocessing, or interpersonal therapy) or did not report symptoms separately for each disorder. To reduce clinical heterogeneity, studies were also excluded if they had done any of the following: (1) used self-guided therapy without any guid-

ance, (2) used CBT combined with medication or pill placebo, or (3) tested inpatients.

Study Selection

Titles and abstracts of the records were independently screened by two of us (E.A.M.vD. and S.C.vV.) with the use of the Covidence systematic review tool.[28] The full-text screening and data extraction were independently performed by two of us (E.A.M.vD. and R.M.vdH.). In case of disagreements during the screening or data extraction process, a consensus was reached through discussion or by the decision of a third person (P.C.). If full-text records were inaccessible, authors and/or libraries were contacted (k = 12; response rate, 33%). If crucial statistics were missing, study authors were contacted (k = 8; response rate, 38%).

Quality Assessment

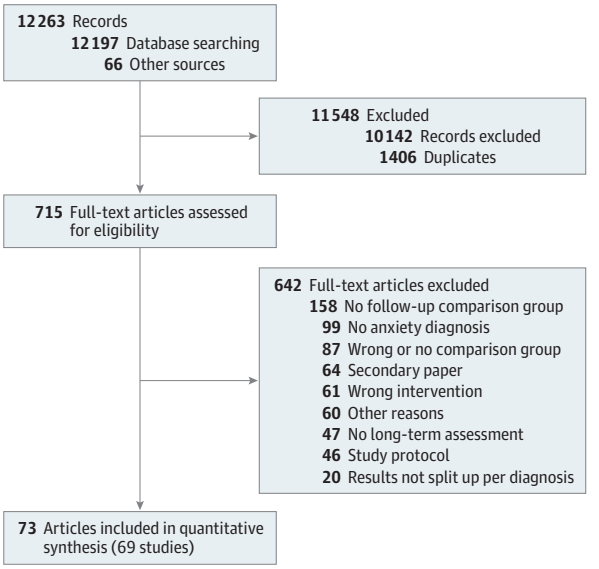
To assess the quality of the included studies, 5 criteria of the Cochrane Collaboration's risk of bias tool were used: adequate generation of allocation sequence, concealment of allocation to conditions, blinding of outcome assessment, adequately dealing with incomplete outcome data (this was evaluated as being of high quality when we could use intention-to-treat analyses), and no selective outcome reporting (based on whether authors referred to trial registrations or study design publications).[29] In addition, quality of treatment implementation was evaluated according to the following 4 criteria outlined by Chambless and Hollon[30]: (1) the use of a treatment protocol, (2) training of therapists, (3) monitoring of therapy (integrity check), and (4) researcher allegiance. Researcher allegiance was defined as 1 of the authors' involvement in developing the treatment under investigation, except when collaborators had mixed allegiances.[31] All quality assessments were independently completed by two of us (E.A.M.vD. and R.M.vdH.), and disagreement was solved through discussion or by the decision of a third person (P.C.).

Statistical Analysis

Comprehensive Meta-analysis software, version 3 (Biostat)[32] was used to calculate the pooled effect sizes separately for each disorder. If studies used multiple symptom measures, these outcomes were pooled within studies,[33] except for a sensitivity analysis that included 1 outcome measure (based on a frequency ranking). Random-effects models were selected in all analyses and available intention-to-treat data were used. Power analyses were conducted with the online Power Calculator Tool.[34] The primary outcome variable was anxiety symptoms. Hedges g was calculated to indicate differences between treatment and comparison groups at posttreatment and follow-up. Follow-up measurements were categorized into 3 periods: 1 to 6 months, 6 to 12 months, and 12 months or more of posttreatment follow-up. Relapse rates were defined as the percentage of relapse after treatment response at follow-up (treatment group vs comparison group). Relative risk was calculated to indicate dropout differences between treatment and comparison groups. Subgroup analyses were performed on treatment approaches, comparison groups, and study quality using a mixed-effects model and meta-regression. Analyses with at least 3 studies per subgroup are reported.

To assess potential publication bias, the Egger test of the intercept was used, which is a significance test based on the asymmetry of funnel plots.[35] The funnel-plot–based method of Duval and Tweedie[36] was used to test and adjust for publication bias through a trim and fill technique. To estimate heterogeneity across studies, the I^2 statistic with 95% CIs (using the HETEROGI module for Stata, version 8 [StataCorp])[37] was calculated, which displays the proportion of the observed variance that would remain if we could remove the sampling error. A common benchmark for interpretation is 25% for small, 50% for medium, and 75% for large heterogeneity.[33] We also calculated 95% prediction intervals to estimate the effect size range in future studies.[38]

Figure 1. Flow Diagram of Selection and Inclusion Process

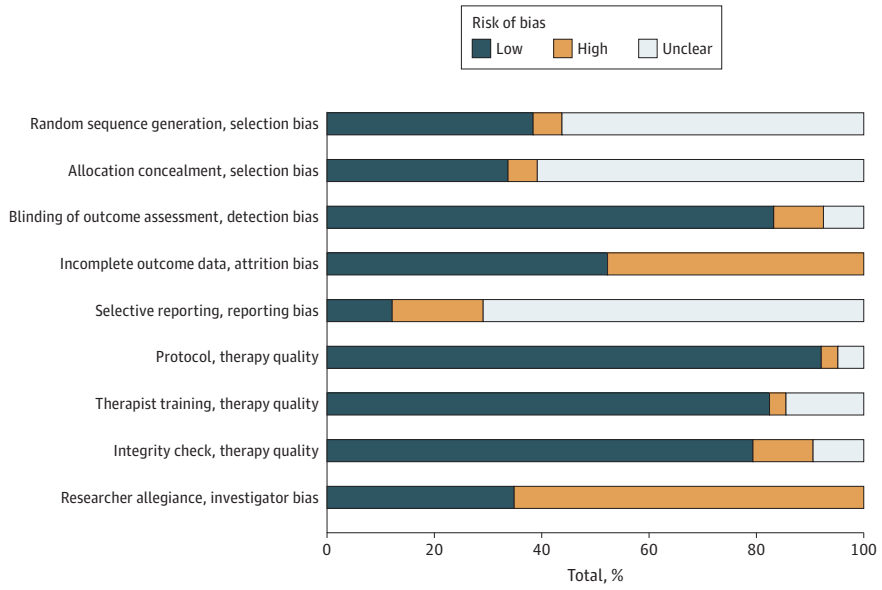


Results

Selection and Characteristics of Included Studies

Figure 1 displays the PRISMA flowchart of the selection and inclusion process. We screened 10 857 titles and abstracts and retrieved 715 full-text records, of which 69 published studies (reported in 73 records) met our inclusion criteria: 14 studies on GAD, 13 studies on PD, 7 studies on SAD, 3 studies on specific phobia, 30 studies on PTSD, and 2 studies on OCD (eTable 2 in the Supplement presents characteristics of these studies). A total of 4118 unique patients were enrolled (age and sex not available in the final analyses). The studies examined CBT (number of studies [k] = 42), exposure therapy, (k = 26), cognitive therapy (k = 10), cognitive reprocessing (k = 1), metacognitive therapy (k = 1), applied tension (k = 1), and acceptance and commitment therapy (k = 1). Comparison groups consisted of care as usual (k = 13), relaxation (k = 24), psychoeducation (k = 2), pill placebo (k = 5), supportive therapy

Figure 2. Study Design Quality, Therapy Quality, and Researcher Allegiance



(k = 14), waiting list (k = 12), and tension only (k = 1). Multiple treatment or comparison groups within 1 study were pooled together (k = 9). We found 41 studies reporting outcomes at 1 to 6 months, 34 studies at 6 to 12 months, and 24 studies at 12 months or more of follow-up. Groups did generally not differ in dropout (relative risk range, 0.97-1.03; P > .50), but for PTSD, there was slightly more dropout in the comparison group (relative risk, 0.95; P = .01).

Quality Assessments

Figure 2 and eFigure 1 in the Supplement present the study and treatment quality assessments. Only 12 studies met criteria for high quality (ie, ≥4 of 5 criteria). Nineteen of the studies (27.5%) applied random sequence generation and allocation concealment. In 44 studies (63.8%), the outcome assessments were blinded and 35 studies (50.7%) applied intention-to-treat analyses. Only 21 studies (30.4%) reported a preregistration or a design protocol, and in 13 cases, the outcomes were not reported in accordance with their preregistration. The overall treatment implementation quality was high and most studies had a high risk of researchers' allegiance.

Main Analyses

Table 1 presents effect sizes, heterogeneity indices, and adjusted effect sizes for risk of publication bias based on the trim and fill procedure of Duval and Tweedie[36] for all disorders across time (eFigures 2-7 in the Supplement provide forest plots and eFigures 8-12 in the Supplement provide funnel plots). A sensitivity analysis with 1 outcome measure yielded similar results (eTable 3 in the Supplement). After treatment, the pooled effect size of CBT relative to control conditions was small for PD (Hedges g, 0.22; 95% CI, 0.01-0.43); medium for GAD (Hedges g, 0.39; 95% CI, 0.12-0.66), SAD (Hedges g, 0.38; 95% CI, 0.19-0.57), and specific phobia (Hedges g, 0.49; 95% CI, 0.13-0.84); and medium to large for PTSD (Hedges g, 0.72; 95% CI, 0.52-0.93) and OCD (Hedges g, 0.70; 95% CI, 0.29-1.12). The Egger test of the intercept was only significant for PTSD (intercept β, 3.13; 95% CI, 1.78-4.49, P < .01; all others, β < 2.34; P > .20). The trim and fill procedure[36] yielded lower adjusted effect sizes for all disorders except OCD (Table 1). Heterogeneity was low to moderate for PD, SAD, specific phobia, and OCD, and it was moderate to large for GAD and PTSD.

At 1 to 6 months of follow-up, the relative pooled estimate of CBT was small for GAD (Hedges g, 0.07; 95% CI, −0.50 to 0.63) and PD (Hedges g, 0.27; 95% CI, −0.01 to 0.55), medium for SAD (Hedges g, 0.60; 95% CI, 0.36-0.85), and medium to large for specific phobia (Hedges g, 0.72; 95% CI, 0.01-1.44), PTSD (Hedges g, 0.67; 95% CI, 0.46-0.88), and OCD (Hedges g, 0.85; 95% CI, 0.47-1.22). The Egger test of the intercept was significant for GAD (intercept β, −10.45; 95% CI, −16.15 to 4.76, P = .03) and PTSD (intercept β, 3.10; 95% CI, 1.28-4.92, P = .002; all others: β < 4.22, P > .08), and the trim and fill procedure resulted in a lower adjusted effect size only for PTSD (Hedges g, 0.50; 95% CI, 0.27-0.73). Heterogeneity was low for PD, SAD, and OCD; moderate for specific phobia; and moderate to large for GAD and PTSD.

At 6 to 12 months of follow-up, the pooled effect size of CBT relative to control conditions was small to medium for GAD (Hedges g, 0.40; 95% CI, 0.13-0.67), PD (Hedges g, 0.35; 95% CI, 0.11-0.59), and SAD (Hedges g, 0.34; 95% CI, 0.07-0.61) and medium for PTSD (Hedges g, 0.59; 95% CI, 0.42-0.77). No pooled effect sizes could be calculated for specific phobia (k = 0) and OCD (k = 0). The Egger test of the intercept did not indicate a risk of publication bias for any disorder (all β < 2.74, P > .06). The trim and fill procedure resulted in lower adjusted effect sizes only for SAD and PTSD (Table 1). Heterogeneity was low for PD, SAD, and PTSD and moderate for GAD.

After a follow-up of 12 months or more, CBT was still associated with a better outcome than control conditions for GAD

Table 1. Treatment Effects (Hedges *g*), Heterogeneity Indices, and Effect Sizes Adjusted for Publication Bias Across Time

Diagnosis	*k*	Hedges *g* (95% CI)	95% Prediction Interval	(95% CI) I^2	Adjusted Hedges *g*
At Treatment Completion					
GAD	14	0.39 (0.12 to 0.66)[a]	−0.55 to 1.33	67 (42-81)	0.34 (0.05 to 0.62)[a]
PD	13	0.22 (0.01 to 0.43)[a]	−0.30 to 0.74	29 (0-63)	0.19 (−0.02 to 0.41)
SAD	7	0.38 (0.19 to 0.57)[a]	0.04 to 0.72	11 (0-63)	0.22 (−0.01 to 0.44)
Specific phobia	3	0.49 (0.13 to 0.84)[a]	−1.80 to 2.78	0 (0-90)	0.34 (0.04 to 0.63)[a]
PTSD	30	0.72 (0.52 to 0.93)[a]	−0.26 to 1.71	74 (62-81)	0.50 (0.28 to 0.72)[a]
OCD	2	0.70 (0.29 to 1.12)[a]	NA	17 (NA)	NA
1-6 mo of Follow-up					
GAD	3	0.07 (−0.50 to 0.63)[b]	−6.48 to 6.61	73 (10-92)	NA[c]
PD	6	0.27 (−0.01 to 0.55)	−0.22 to 0.76	8 (0-64)	NA[c]
SAD	4	0.60 (0.36 to 0.85)[a]	0.06 to 1.15	0 (0-68)	NA[c]
Specific phobia	2	0.72 (0.01 to 1.44)[a]	NA	39 (NA)	NA
PTSD	24	0.67 (0.46 to 0.88)[a]	−0.19 to 1.52	63 (38-75)	0.50 (0.27 to 0.73)[a]
OCD	2	0.85 (0.47 to 1.22)[a]	NA	0 (NA)	NA
6-12 mo of Follow-up					
GAD	11	0.40 (0.13 to 0.67)[a]	−0.41 to 1.22	59 (20-79)	NA[c]
PD	9	0.35 (0.11 to 0.59)[a]	−0.08 to 0.77	12 (0-60)	NA[c]
SAD	3	0.34 (0.07 to 0.61)[a]	−1.40 to 2.08	0 (0-73)	0.22 (0.01 to 0.45)[a]
Specific phobia	0	NA	NA	NA	NA
PTSD	11	0.59 (0.42 to 0.77)[a]	0.28 to 0.90	12 (0-57)	0.55 (0.35 to 0.75)[a]
OCD	0	NA	NA	NA	NA
≥12 mo of Follow-up					
GAD	10	0.22 (0.02 to 0.42)[a]	−0.18 to 0.61	18 (0-59)	NA[c]
PD	5	0.14 (−0.19 to 0.47)[b]	−0.40 to 0.67	0 (0-64)	NA[c]
SAD	3	0.42 (0.04 to 0.79)[a]	−2.00 to 2.83	0 (0-73)	NA[c]
Specific phobia	1	NA	NA	NA	NA
PTSD	5	0.84 (0.03 to 1.64)[a]	−2.13 to 3.80	88 (71-93)	0.54 (−0.20 to 1.29)
OCD	0	NA	NA	NA	NA

Abbreviations: GAD, generalized anxiety disorder; *k*, number of studies; NA, not applicable; OCD, obsessive-compulsive disorder; PD, panic disorder with or without agoraphobia; PTSD, posttraumatic stress disorder; SAD, social anxiety disorder.

[a] Effect sizes that are statistically significant ($P < .05$).

[b] Post hoc statistical power beneath 80% ($\alpha = .05$).

[c] No adjustment for publication bias based on the trim and fill procedure of Duval and Tweedie.[36]

(Hedges *g*, 0.22; 95% CI, 0.02-0.42; $k = 10$), SAD (Hedges *g*, 0.42; 95% CI, 0.04-0.79; $k = 3$), and PTSD (Hedges *g*, 0.84; 95% CI, 0.03-1.64; $k = 5$), but this effect was not significant for PD (Hedges *g*, 0.14; 95% CI, -0.19 to 0.47; $k = 5$) and could not be calculated for specific phobia ($k = 1$) and OCD ($k = 0$). The Egger test of the intercept did not indicate a risk of publication bias ($\beta < 3.51$ for all, $P > .09$), but the trim and fill procedure yielded a lower nonsignificant effect for PTSD (Hedges *g*, 0.54; 95% CI, -0.20 to 1.29). Heterogeneity was low for PD, SAD, and GAD but large for PTSD.

Subgroup Analyses

eTables 4 and 5 in the Supplement present exploratory subgroup analyses for treatment approaches and comparison groups. For specific phobia and OCD, subgroup analyses could not be performed (<2 studies per comparison group). Meta-regression analyses revealed no significant differences across treatment approaches for any disorder at any time (all $Q < 1.92$; $P > .38$).

For GAD and SAD, the comparison groups did not significantly differ at any time. For PD, subgroup analyses showed a significant medium treatment effect of CBT relative to pill placebo at posttreatment (Hedges *g*, 0.42) and at 6 to 12 months of follow-up (Hedges *g*, 0.73). There were no significant treatment effects relative to any other active comparison group at any time (all $P > .06$; eTable 5 in the Supplement). For PTSD, CBT appeared to be generally more effective relative to all comparison groups until 12 months of follow-up (Hedges *g*, >0.73; all $P < .02$), but not compared with supportive therapy after 12 months or more (Hedges *g*, 0.08; $P = .44$). At treatment completion, studies that used a waiting list comparison group yielded significantly ($P < .01$) larger effect sizes (Hedges *g*, 1.25), while studies using a supportive therapy comparison condition yielded significantly lower effect sizes (Hedges *g*, 0.27) ($P = .02$).

Exploratory subgroup analyses on study quality could only be performed for PTSD (high-quality studies: $k = 8$) and showed larger effect sizes at all times for high-quality studies (Hedges *g*, 0.65-2.10) compared with the other studies (Hedges *g*, 0.51-0.57). There were no high-quality studies for SAD and specific phobia and only a few for PD ($k = 1$), GAD ($k = 2$), and OCD ($k = 1$).

Table 2. Number of Relapses After Successful CBT

Source	Diagnosis	Instrument	Criterion		Follow-up, mo	No. of Relapses/Total No.	
			Responder	Relapse		Treatment	Control
Arntz and van den Hout,[43] 1996	PD	Diary	No panic attack in 2 wk	Pretest panic attack frequency	7	CT, 2/14	Applied relaxation, 0/9
Barlow et al,[42] 2000	PD	PDSS	40% Reduction from baseline on PDSS	Not meeting responder criterion	12	CBT, 1/24	Pill placebo, 0/3
Öst et al,[39] 1993	PD	Percentage of BAT completed	Clinically significant improvement on BAT[46]	Not meeting responder criterion	12	Exposure, 0/12 CT 0/9	Applied relaxation, 0/13
Öst and Westling,[40] 1995	PD	Diary	No panic attack in 3 wk	Not meeting responder criterion	12	CBT, 0/14	Applied relaxation, 0/11
Shear et al,[41] 2001	PD	CGI	CGI improvement: >2 (much improved) and CGI severity: <3 (mild)	Not meeting responder criterion	6	CBT, 0/16	Pill placebo, 0/3
Simpson et al,[44] 2004[a]	OCD	CGI	CGI improvement: much improved relative to week 0	(1) A return to pretreatment severity or worse in the past week on the CGI Severity subscale or (2) An unsafe clinical state based on the clinical judgment of the treating clinician	3	EX/RP, 2/18	Pill placebo, 0/1

Abbreviations: BAT, behavioral avoidance test (based on agoraphobic situations hierarchy); CBT, cognitive behavioral therapy; CGI, Clinical Global Impression Scale; CT, cognitive therapy; EX/RP, exposure with response prevention; OCD, obsessive-compulsive disorder; PD, panic disorder with or without agoraphobia; PDSS, Panic Disorder Severity Scale.

[a] This study was not included in the meta-analysis; it only examined treatment responders at follow-up.

Relapse

A total of 6 studies (7 comparisons) reported relapse rates after successful treatment. Of these, 5 studies were about PD[39-43] and 1 was about OCD.[44] An additional study described relapse of PD as a comorbid condition after PTSD treatment, and this study was not included.[45] All 6 studies used small sample sizes (n < 28), and most operationalized successful treatment using ambiguous treatment response criteria rather than reliable remission criteria (eg, the absence of a disorder based on a clinical interview). Therefore, we refrained from statistically pooling these results and instead presented outcomes per study in **Table 2**. Overall, relapse rates were relatively low: in 3 of 7 comparisons, relapse occurred after successful CBT and relapse rates ranged from 0% to 14%.

Discussion

Summary of Results

This systematic review and meta-analysis examined the long-term outcome of CBT for anxiety disorders, PTSD, and OCD across 69 randomized clinical trials. Overall, CBT was associated with moderate symptom reductions up to 12 months after treatment. Longer effects were still significant for GAD, SAD, and PTSD, but not for PD and could not be calculated for specific phobia and OCD. Because this meta-analysis included a limited number of high-quality studies and English language articles only, our reported effect estimates should be interpreted with caution. Because statistical heterogeneity was considerable in GAD and PTSD studies, our effect estimates for these disorders are uncertain. Future meta-analyses should aim to explain this heterogeneity as more studies become available. Although post hoc power analyses generally showed sufficient statistical power of our main analyses, simulation studies showed that at least 40 studies per analysis are needed to reach sufficient power.[47] Therefore, nonsignificant findings, especially of the subgroup analyses, should be interpreted as the absence of evidence rather than evidence of absence.

Our overall findings were in line with CBT outcomes for depression[48] and suggest that skills and insights acquired during CBT are relatively stable until 12 months after treatment but do not improve further. Nevertheless, evidence for CBT outcomes at 12 months or more after treatment is scarce. Given the chronic trajectories of anxiety-related disorders[17] and because longer illness duration may increase the odds of developing comorbidity,[49] it is important to examine whether treatment effects are maintained 12 months or more after treatment. Thus, more research on CBT efficacy at 12 months or more of follow-up and on ways to optimize effects is needed.

Relapse rates after successful CBT were relatively low (0%-14%) compared with uncontrolled trials that indicated a maximum relapse of 13% for SAD[50] and 23% for PD.[51] However, only a few studies reported them (5 studies for PD and 1 for OCD), in contrast to studies on pharmacotherapy for anxiety-related disorders that frequently report clinical relapse after treatment discontinuation.[52] Also, these studies calculated relapse rates based on ambiguous response criteria rather than relative to complete remission. Therefore, future research should carefully define and report relapse criteria (eg, a return of the full symptomatology[24,53] based on a structured interview). Future research may also give insight into risk factors for relapse, which could identify patients at risk who may benefit from additional or more intensive therapy or from pharmacotherapy to prevent relapse. Relapse prevention after psy-

chotherapy is still relatively uncharted in the field of anxiety-related disorders but is quite common and effective in depressive disorders.[54] For example, studies have shown the efficacy of well-being therapy[55,56] as second-line relapse prevention strategy in patients with GAD.[57]

For PD, when corrected for publication bias, CBT outcome did not significantly differ from control conditions (except for a small to medium effect at 6-12 months of follow-up). This may be explained by the frequent use of applied relaxation as a control condition, which may involve some exposure.[39] Relaxation appeared to be as effective as CBT in a previous meta-analysis.[20] Subgroup analyses across comparison groups revealed a medium treatment effect for PD within 12 months after treatment when CBT was compared with pill placebo, but not relative to other active comparison groups. However, the subgroup analyses should be interpreted with caution because of the small subsample sizes.

For specific phobia and OCD, only a few studies met our inclusion criteria, and treatment effect estimates could not be calculated beyond a 6-month follow-up. Most previous studies on OCD treatment with long-term assessments have tested the efficacy of pharmacotherapy (augmented with CBT).[58,59] Because approximately 50% of patients with OCD do not respond to pharmacotherapy and many patients relapse after medication discontinuation,[58] more research is needed on the long-term efficacy of CBT as an alternative stand-alone treatment.

Regarding PTSD, after correcting for publication bias, we observed medium treatment effects favoring CBT over control conditions at posttreatment until 12 months of follow-up. At 12 months or more of follow-up, there was a nonsignificant medium effect adjusted for publication bias, which probably did not reach statistical significance because of limited statistical power.

Strengths and Limitations

Strengths of this meta-analysis are the inclusion of more comparison groups, which yielded more studies than previous meta-analyses,[18-21] and the investigation of long-term outcomes (including relapse rates) after CBT for anxiety-related disorders. Furthermore, we conducted a comprehensive literature search, an independent screening and data extraction, and treatment and study quality assessments. Several limitations should also be noted. First, meta-analyses are inherently associated with heterogeneity regarding methodological aspects (eg, outcome measures) and clinical aspects (eg, CBT approaches and samples). Therefore, future research is needed to test which specific methodological or treatment factors explain the reported effects.[60] Second, because of limited experimental control during follow-up periods, confounding factors may have threatened the validity of our long-term effect estimates (eg, because of additional treatment or adverse life events). Third, symptom outcome measures were averaged to handle dependent outcomes, which may have resulted in overestimated SEs.[61] Fourth, most studies had suboptimal designs (or these criteria were poorly reported) and a high risk of researcher allegiance bias, which may have affected the reliability of our effect estimates.

Conclusions

Anxiety-related disorders are characterized by a chronic course, thus sustainable treatment effects are important. The results of this meta-analysis suggest that, on average, CBT was associated with moderate symptom reductions in anxiety disorders, PTSD, and OCD until 12 months after treatment completion. At a follow-up of 12 months or more, these effects were still present for GAD, SAD, and PTSD, but not for PD. For specific phobia and OCD, no follow-up data beyond 6 months after treatment completion were available. Studies on relapse were scarce but gave the preliminary impression that relapse rates after successful treatment, predominantly for PD, may be relatively low (0%-14% at 3-12 months following treatment completion). More high-quality randomized clinical trials on long-term treatment effects (preferably ≥12 months after treatment completion) and relapse are warranted to facilitate more reliable long-term effect size estimations.

ARTICLE INFORMATION

Accepted for Publication: September 18, 2019.

Published Online: November 23, 2019. doi:10.1001/jamapsychiatry.2019.3986

Correction: This article was corrected on May 13, 2020, to fix the description of the last category of follow-up time throughout the Abstract, the text, Table 1, and the Supplement.

Author Contributions: Ms van Dis had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: van Dis, van Veen, Hagenaars, Batelaan, Bockting, Cuijpers, Engelhard.
Acquisition, analysis, or interpretation of data: van Dis, van Veen, Hagenaars, Bockting, van den Heuvel.
Drafting of the manuscript: van Dis, Engelhard.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: van Dis.
Obtained funding: Bockting, Engelhard.
Administrative, technical, or material support: van Dis, Engelhard.
Supervision: Hagenaars, Batelaan, Bockting, Cuijpers, Engelhard.
Full-text screening, data extraction, quality assessments of articles: van den Heuvel.

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was supported by Vici grant 453-15-005 from the Netherlands Organization for Scientific Research (Dr Engelhard). Ms van Veen was supported by TOP grant 40-00812-98-12030 (Dr Engelhard) from the Netherlands Organization for Health Research and Development. Dr Bockting was supported by the fellowship of the Institute for Advanced Studies of the University of Amsterdam.

Role of the Funder/Sponsor: The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Meeting Presentation: This paper was presented at the Association for Cognitive and Behavioral Therapies meeting; November 23, 2019; Atlanta, Georgia.

Additional Contributions: We thank Paulien Wiersma, MSc (Utrecht University Library), for her assistance in designing the search strategy.

REFERENCES

1. Baxter AJ, Scott KM, Vos T, Whiteford HA. Global prevalence of anxiety disorders: a systematic review and meta-regression. *Psychol Med*. 2013;43(5):897-910. doi:10.1017/S003329171200147X

2. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication [published

correction appears in *Arch Gen Psychiatry*. 2005;62(7):709]. *Arch Gen Psychiatry*. 2005;62(6):617-627. doi:10.1001/archpsyc.62.6.617

3. Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382(9904):1575-1586. doi:10.1016/S0140-6736(13)61611-6

4. Smit F, Cuijpers P, Oostenbrink J, Batelaan N, de Graaf R, Beekman A. Costs of nine common mental disorders: implications for curative and preventive psychiatry. *J Ment Health Policy Econ*. 2006;9(4):193-200.

5. Greenberg PE, Sisitsky T, Kessler RC, et al. The economic burden of anxiety disorders in the 1990s. *J Clin Psychiatry*. 1999;60(7):427-435. doi:10.4088/JCP.v60n0702

6. Layard R, Clark DM. *Thrive: How Better Mental Health Care Transforms Lives and Saves Money*. Princeton, NJ: Princeton University Press; 2015.

7. Stein MB, Goin MK, Pollack MH, et al. *Practice Guideline for the Treatment of Patients With Panic Disorder*. 2nd ed. Washington, DC: American Psychiatric Press; 2009, doi:10.1136/bmj.1.4496.313-d

8. Benedek DM, Friedman MJ, Zatzick D, Ursano RJ. *Guideline Watch (March 2009): Practice Guideline for the Treatment of Patients With Acute Stress Disorder and Postttraumatic Stress Disorder*. Arlington, VA: American Psychiatric Association; 2009. doi:10.1176/appi.books.9780890423363.149073

9. Koran LM, Simpson HB. Guideline Watch (March 2013). Practice Guideline for the Treatment of Patients with Obsessive-Compulsive Disorder. Arlington, VA: American Psychiatric Association; 2013. https://www.psychiatry.org/psychiatrists/practice/clinical-practice-guidelines. Accessed November 12, 2019.

10. National Institute for Health and Care Excellence. Post-traumatic stress disorder. Clinical guideline NG116. Published 2018. https://www.nice.org.uk/guidance/ng116/. Accessed February 13, 2019.

11. National Institute for Health and Care Excellence. Generalised anxiety disorder and panic disorder in adults: management. Clinical guideline CG113. Published 2011. https://www.nice.org.uk/guidance/cg113. Accessed February 13, 2019.

12. National Institute for Health and Care Excellence. Social anxiety disorder: recognition, assessment and treatment. Clinical guideline CG159. Published 2013. https://www.nice.org.uk/guidance/cg159. Accessed February 13, 2019.

13. National Institute for Health and Care Excellence. Obsessive-compulsive disorder and body dysmorphic disorder: treatment. Clinical guideline CG31. Published 2005. https://www.nice.org.uk/guidance/cg31. Accessed February 13, 2019.

14. McHugh RK, Whitton SW, Peckham AD, Welge JA, Otto MW. Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review. *J Clin Psychiatry*. 2013;74(6):595-602. doi:10.4088/JCP.12r07757

15. Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Huibers MJH. How effective are cognitive behavior therapies for major depression and anxiety disorders? a meta-analytic update of the evidence. *World Psychiatry*. 2016;15(3):245-258. doi:10.1002/wps.20346

16. Hofmann SG, Smits JA. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. *J Clin Psychiatry*. 2008;69(4):621-632. doi:10.4088/JCP.v69n0415

17. Hofmeijer-Sevink MK, Batelaan NM, van Megen HJ, et al. Clinical relevance of comorbidity in anxiety disorders: a report from the Netherlands Study of Depression and Anxiety (NESDA). *J Affect Disord*. 2012;137(1-3):106-112. doi:10.1016/j.jad.2011.12.008

18. Bandelow B, Sagebiel A, Belz M, Görlich Y, Michaelis S, Wedekind D. Enduring effects of psychological treatments for anxiety disorders: meta-analysis of follow-up studies. *Br J Psychiatry*. 2018;212(6):333-338. doi:10.1192/bjp.2018.49

19. Carpenter JK, Andrews LA, Witcraft SM, Powers MB, Smits JAJ, Hofmann SG. Cognitive behavioral therapy for anxiety and related disorders: a meta-analysis of randomized placebo-controlled trials. *Depress Anxiety*. 2018;35(6):502-514. doi:10.1002/da.22728

20. Montero-Marin J, Garcia-Campayo J, López-Montoyo A, Zabaleta-del-Olmo E, Cuijpers P. Is cognitive–behavioural therapy more effective than relaxation therapy in the treatment of anxiety disorders? a meta-analysis. *Psychol Med*. 2017;48(9):1427-1436. doi:10.1017/S0033291717003099

21. Springer KS, Levy HC, Tolin DF. Remission in CBT for adult anxiety disorders: a meta-analysis. *Clin Psychol Rev*. 2018;61:1-8. doi:10.1016/j.cpr.2018.03.002

22. Cuijpers P, Weitz E, Cristea IA, Twisk J. Pre-post effect sizes should be avoided in meta-analyses. *Epidemiol Psychiatr Sci*. 2017;26(4):364-368. doi:10.1017/S2045796016000809

23. Scholten WD, Batelaan NM, Penninx BW, et al. Diagnostic instability of recurrence and the impact on recurrence rates in depressive and anxiety disorders. *J Affect Disord*. 2016;195:185-190. doi:10.1016/j.jad.2016.02.025

24. Vervliet B, Craske MG, Hermans D. Fear extinction and relapse: state of the art. *Annu Rev Clin Psychol*. 2013;9(1):215-248. doi:10.1146/annurev-clinpsy-050212-185542

25. PROPERO. Long-term Effects of Psychotherapy for Anxiety Disorders: A Meta-analysis. CRD42017067363. https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=67363. Accessed November 9, 2019.

26. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264-269. doi:10.7326/0003-4819-151-4-200908180-00135

27. Loerinc AG, Meuret AE, Twohig MP, Rosenfield D, Bluett EJ, Craske MG. Response rates for CBT for anxiety disorders: need for standardized criteria. *Clin Psychol Rev*. 2015;42:72-82. doi:10.1016/j.cpr.2015.08.004

28. Covidence. http://www.covidence.org. Accessed Accessed November 9, 2019.

29. Higgins JPT, Altman DG, Sterne JAC, eds. Chapter 8: assessing risk of bias in included studies. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. London, England: Cochrane Collaboration; 2011. https://handbook-5-1.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm. Updated March 2011. Accessed November 12, 2019.

30. Chambless DL, Hollon SD. Defining empirically supported therapies. *J Consult Clin Psychol*. 1998;66(1):7-18. doi:10.1037/0022-006X.66.1.7

31. Munder T, Brütsch O, Leonhart R, Gerger H, Barth J. Researcher allegiance in psychotherapy outcome research: an overview of reviews. *Clin Psychol Rev*. 2013;33(4):501-511. doi:10.1016/j.cpr.2013.02.002

32. Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive Meta-analysis, Version 3.0. Biostat Inc. https://www.meta-analysis.com/?gclid=EAIaIQobChMI3vatuP3k5QIVqx-tBh2MwQDHEAAYASAAEgLSGfD_BwE. Published 2013. Accessed November 12, 2019.

33. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-analysis*. Chichester, England: John Wiley & Sons Ltd; 2009. doi:10.1002/9780470743386

34. Harrer M, Cuijpers P, Furukawa TA, Ebert DD. Doing meta-analysis in R: a hands-on guide. Zenodo website. https://zenodo.org/record/2551803#.Xccy8697nHY. Published 2019. Accessed November 9, 2019.

35. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629

36. 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. doi:10.1111/j.0006-341X.2000.00455.x

37. Orsini N, Bottai M, Higgins J, Buchan I. HETEROGI: Stata module to quantify heterogeneity in a meta-analysis. https://econpapers.repec.org/software/bocbocode/s449201.htm. Published 2006. Accessed March 21, 2019.

38. Borenstein M, Higgins JP, Hedges LV, Rothstein HR. Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Res Synth Methods*. 2017;8(1):5-18. doi:10.1002/jrsm.1230

39. Öst LG, Westling BE, Hellström K. Applied relaxation, exposure in vivo and cognitive methods in the treatment of panic disorder with agoraphobia. *Behav Res Ther*. 1993;31(4):383-394. doi:10.1016/0005-7967(93)90095-C

40. Öst LG, Westling BE. Applied relaxation vs cognitive behavior therapy in the treatment of panic disorder. *Behav Res Ther*. 1995;33(2):145-158. doi:10.1016/0005-7967(94)E0026-F

41. Shear MK, Houck P, Greeno C, Masters S. Emotion-focused psychotherapy for patients with panic disorder. *Am J Psychiatry*. 2001;158(12):1993-1998. doi:10.1176/appi.ajp.158.12.1993

42. Barlow DH, Gorman JM, Shear MK, Woods SW. Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: a randomized controlled trial. *JAMA*. 2000;283(19):2529-2536. doi:10.1001/jama.283.19.2529

43. Arntz A, van den Hout M. Psychological treatments of panic disorder without agoraphobia: cognitive therapy versus applied relaxation. *Behav Res Ther*. 1996;34(2):113-121. doi:10.1016/0005-7967(95)00061-5

44. Simpson HB, Liebowitz MR, Foa EB, et al. Post-treatment effects of exposure therapy and

clomipramine in obsessive-compulsive disorder. *Depress Anxiety*. 2004;19(4):225-233. doi:10.1002/da.20003

45. Vaughan K, Armstrong MS, Gold R, O'Connor N, Jenneke W, Tarrier N. A trial of eye movement desensitization compared to image habituation training and applied muscle relaxation in post-traumatic stress disorder. *J Behav Ther Exp Psychiatry*. 1994;25(4):283-291. doi:10.1016/0005-7916(94)90036-1

46. Jacobson NS, Follette WC, Revenstorf D. Psychotherapy outcome research: methods for reporting variability and evaluating clinical significance. *Behav Ther*. 1984;15(4):336-352. doi:10.1016/S0005-7894(84)80002-7

47. López-López JA, Marín-Martínez F, Sánchez-Meca J, Van den Noortgate W, Viechtbauer W. Estimation of the predictive power of the model in mixed-effects meta-regression: a simulation study. *Br J Math Stat Psychol*. 2014;67(1):30-48. doi:10.1111/bmsp.12002

48. Cuijpers P, Hollon SD, van Straten A, Bockting C, Berking M, Andersson G. Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? a meta-analysis. *BMJ Open*. 2013;3(4):e002542. doi:10.1136/bmjopen-2012-002542

49. van Oudheusden LJB, Eikelenboom M, van Megen HJGM, et al. Chronic obsessive-compulsive disorder: prognostic factors. *Psychol Med*. 2018;48(13):2213-2222. doi:10.1017/S0033291717003701

50. Fava GA, Grandi S, Rafanelli C, Ruini C, Conti S, Belluardo P. Long-term outcome of social phobia treated by exposure. *Psychol Med*. 2001;31(5):899-905. doi:10.1017/S0033291701004020

51. Fava GA, Rafanelli C, Grandi S, et al. Long-term outcome of panic disorder with agoraphobia treated by exposure. *Psychol Med*. 2001;31(5):891-898. doi:10.1017/S0033291701003592

52. Batelaan NM, Bosman RC, Muntingh A, Scholten WD, Huijbregts KM, van Balkom AJLM. Risk of relapse after antidepressant discontinuation in anxiety disorders, obsessive-compulsive disorder, and post-traumatic stress disorder: systematic review and meta-analysis of relapse prevention trials. *BMJ*. 2017;358:j3927. doi:10.1136/bmj.j3927

53. Rachman S. The return of fear: review and prospect. *Clin Psychol Rev*. 1989;9(2):147-168. doi:10.1016/0272-7358(89)90025-1

54. Bockting CL, Hollon SD, Jarrett RB, Kuyken W, Dobson K. A lifetime approach to major depressive disorder: the contributions of psychological interventions in preventing relapse and recurrence. *Clin Psychol Rev*. 2015;41:16-26. doi:10.1016/j.cpr.2015.02.003

55. Fava GA, Rafanelli C, Cazzaro M, Conti S, Grandi S. Well-being therapy: a novel psychotherapeutic approach for residual symptoms of affective disorders. *Psychol Med*. 1998;28(2):475-480. doi:10.1017/S0033291797006363

56. Ruini C, Fava GA. Well-being therapy for generalized anxiety disorder. *J Clin Psychol*. 2009;65(5):510-519. doi:10.1002/jclp.20592

57. Fava GA, Ruini C, Rafanelli C, et al. Well-being therapy of generalized anxiety disorder. *Psychother Psychosom*. 2005;74(1):26-30. doi:10.1159/000082023

58. Fineberg NA, Brown A, Pampaloni I. Evidence-based pharmacotherapy of obsessive–compulsive disorder. In: Stein D, Lerer B, Stahl S, eds. *Essential Evidence-Based Psychopharmacology*. Cambridge, England: Cambridge University Press; 2012:128-170. doi:10.1017/CBO9780511910395.009

59. Sharma E, Thennarasu K, Reddy YCJ. Long-term outcome of obsessive-compulsive disorder in adults: a meta-analysis. *J Clin Psychiatry*. 2014;75(9):1019-1027. doi:10.4088/JCP.13r08849

60. Cuijpers P, Reijnders M, Huibers MJH. The role of common factors in psychotherapy outcomes. *Annu Rev Clin Psychol*. 2019;15:207-231. doi:10.1146/annurev-clinpsy-050718-095424

61. Moeyaert M, Ugille M, Beretvas SN, Ferron J, Bunuan R, Van den Noortgate W. Methods for dealing with multiple outcomes in meta-analysis: a comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *Int J Soc Res Methodol*. 2017;20(6):559-572. doi:10.1080/13645579.2016.1252189