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.
A
xiety disorders, posttraumatic stress disorder (PTSD), and obsessive-compulsive disorder (OCD) are highly prevalent\textsuperscript{1,2} and are associated with substantial personal\textsuperscript{3} and societal costs.\textsuperscript{4-6} Clinical practice guidelines recommend psychological and pharmacological interventions for anxiety-related disorders,\textsuperscript{7-13} but most patients favor psychotherapy over pharmacotherapy.\textsuperscript{14} Cognitive behavioral therapy (CBT) for these disorders has been associated with reduced symptoms at short term,\textsuperscript{15,16} with small to medium effect sizes adjusted for publication bias and when studies with waiting list comparisons were not taken into account.\textsuperscript{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.\textsuperscript{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.\textsuperscript{18-21} However, in 2 of these,\textsuperscript{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.\textsuperscript{22} The other 2 meta-analyses did use control conditions, but these were limited to placebo,\textsuperscript{19} resulting in 23 studies, or relaxation,\textsuperscript{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.\textsuperscript{23} Research on relapse and the return of fear has become a major focus of fundamental fear and anxiety research,\textsuperscript{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.

**Methods**

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

**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.

**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\textsuperscript{*}) 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 electronic screening of 4 relevant meta-analyses\textsuperscript{18-21} and 1 systematic review.\textsuperscript{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),\textsuperscript{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),\textsuperscript{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.\(^3\) 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).\(^2\) In addition, quality of treatment implementation was evaluated according to the following 4 criteria outlined by Chambless and Hollon\(^3\): (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.\(^3\) 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)\(^3\) was used to calculate the pooled effect sizes separately for each disorder. If studies used multiple symptom measures, these outcomes were pooled within studies,\(^3\) 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.\(^3\) 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.\(^3\) The funnel-plot–based method of Duval and Tweedie\(^3\) 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 HETERO module for Stata, version 8 [StataCorp])\(^3\) 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.\(^3\) We also calculated 95% prediction intervals to estimate the effect size range in future studies.\(^3\)

**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 4,118 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 (\(k = 1\)), meta-cognitive 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...
(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 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 Tweedie36 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 procedure36 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

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

(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 ($β < 3.51$ 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 ($k = 1$).
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 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).
Long-term Outcomes of Cognitive Behavioral Therapy for Anxiety-Related Disorders

Original Investigation Research

Cholecystectomy 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 therapy55,56 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.39 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 403-00812-98-12030 (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


jamapsychiatry.com

© 2019 American Medical Association. All rights reserved.


