Anxiety Screening
Evidence Report and Systematic Review for the US Preventive Services Task Force

Elizabeth A. O'Connor, PhD; Michelle L. Henninger, PhD; Leslie A. Perdue, MPH; Erin L. Coppola, MPH; Rachel G. Thomas, MPH; Bradley N. Gaynes, MD, MPH

IMPORTANCE Anxiety is commonly seen in primary care and associated with substantial burden.

OBJECTIVE To review the benefits and harms of screening and treatment for anxiety and the accuracy of instruments to detect anxiety among primary care patients.

DATA SOURCES MEDLINE, PsychINFO, Cochrane library through September 7, 2022; references of existing reviews; ongoing surveillance for relevant literature through November 25, 2022.

STUDY SELECTION English-language original studies and systematic reviews of screening or treatment compared with control conditions and test accuracy studies of a priori-selected screening instruments were included. Two investigators independently reviewed abstracts and full-text articles for inclusion. Two investigators independently rated study quality.

DATA EXTRACTION AND SYNTHESIS One investigator abstracted data; a second checked accuracy. Meta-analysis results were included from existing systematic reviews where available; meta-analyses were conducted on original research when evidence was sufficient.

MAIN OUTCOMES AND MEASURES Anxiety and depression outcomes; global quality of life and functioning; sensitivity and specificity of screening tools.

RESULTS Of the 59 publications included, 40 were original studies (N = 275489) and 19 were systematic reviews (including ~483 studies [N~81507]). Two screening studies found no benefit for screening for anxiety. Among test accuracy studies, only the Generalized Anxiety Disorder (GAD) GAD-2 and GAD-7 screening instruments were evaluated by more than 1 study. Both screening instruments had adequate accuracy for detecting generalized anxiety disorder (eg, across 3 studies the GAD-7 at a cutoff of 10 had a pooled sensitivity of 0.79 [95% CI, 0.69 to 0.94] and specificity of 0.89 [95% CI, 0.83 to 0.94]). Evidence was limited for other instruments and other anxiety disorders. A large body of evidence supported the benefit of treatment for anxiety. For example, psychological interventions were associated with a small pooled standardized mean difference of ~0.41 in anxiety symptom severity in primary care patients with anxiety (95% CI, −0.58 to −0.23); 10 RCTs [n = 2075]; I^2 = 40.2%); larger effects were found in general adult populations.

CONCLUSIONS AND RELEVANCE Evidence was insufficient to draw conclusions about the benefits or harms of anxiety screening programs. However, clear evidence exists that treatment for anxiety is beneficial, and more limited evidence indicates that some anxiety screening instruments have acceptable accuracy to detect generalized anxiety disorder.
Anxiety symptoms are relatively common among US adults. The 2019 National Household Interview Survey found that 9.5%, 3.4%, and 2.7% of adults had experienced mild, moderate, or severe symptoms of anxiety, respectively, in the past 2 weeks.¹ National data on the current prevalence of anxiety disorders are lacking, but anxiety disorders are associated with impaired quality of life² and functioning³ and substantial economic costs.⁴ One prior review estimated average annual health expenditures attributable to anxiety disorders among countries in the Organization for Economic Cooperation and Development to be $135 billion.⁵ If effective, routine screening could substantially increase the likelihood that patients receive treatment in a timely manner, potentially saving years of distress and reducing economic burden.

This systematic review was conducted to support the US Preventive Services Task Force (USPSTF) in making a recommendation on anxiety screening in adult primary care patients in the US. The USPSTF has never issued a recommendation on screening for anxiety disorders.

Methods

Scope of Review

Figure 1 shows the analytic framework and key questions (KQs) that guided the review, which were developed in consultation with members of the USPSTF and covered screening for depression, anxiety, and suicide risk. There were no deviations from the original research plan. The current publication discusses the evidence on the benefits and harms of screening for and treatment of anxiety disorders in adults, and the accuracy of screening tools. Detailed methods and results are available in the full evidence review.⁷ In addition to addressing the KQs, the full evidence report also discusses contextual questions and includes an appendix addressing what is known about inequities in the etiology or risk factors for mental health conditions, as well as in diagnosis, treatment access and uptake, and treatment outcomes across racial and ethnic groups. A summary of results related to depression and suicide risk screening is included in a separate publication.⁸

Data Sources and Searches

Ovid MEDLINE, the Cochrane Central Register of Controlled Clinical Trials, the Cochrane Database of Systematic Reviews, and PsycINFO were searched through September 7, 2022. Searches bridged from existing foundational reviews if available or began in 1990 if no relevant foundational review was identified. The search start dates were January of 1990 (KQ1 and KQ3), 2014 (KQ2), and 2015 (KQ4 and KQ5).

Detailed search strategies are listed in the eMethods in the Supplement and were supplemented by examining reference lists of relevant reviews. Article alerts and targeted searches of journals to identify major studies published in the interim that may affect the conclusions or understanding of the evidence and the related USPSTF recommendation were used as part of ongoing surveillance. The last surveillance was conducted on November 25, 2022, and identified no studies affecting review conclusions.

Study Selection

Two investigators independently reviewed titles, abstracts, and full-text articles using predefined eligibility criteria. For KQ1, KQ1a, and KQ3 (benefits and harms of screening), randomized clinical trials (RCTs) of adult primary care patients, including pregnant people, investigating the benefits or harms of screening programs for anxiety were included. Screening programs were defined as efforts to screen all eligible members of a defined group (eg, primary care patients seen at study clinics on specified days), on the presumption that a positive screening result would be acted on clinically. Studies that included additional components beyond screening, such as referral support, training in diagnosis or management, and patient materials, were not excluded. Control groups included participants who either were not screened for anxiety (KQ1) or were screened but whose screening results were not given to their primary care clinician (KQ1a).

For KQ2 (test accuracy), diagnostic accuracy studies of a priori-specified screening tools were included: Generalized Anxiety Disorder scale (GAD), in any form; Patient Health Questionnaire anxiety scale; Edinburgh Perinatal Depression Scale anxiety subscale, for perinatal persons; Geriatric Anxiety Inventory and Geriatric Anxiety Scale for older adults. These tools had been identified as being the most widely used or recommended, based on recommendations of professional societies and government entities, systematic reviews, implementation studies, and clinicians working in some large health systems.

For KQ4 and KQ5 (benefits and harms of treatment), RCTs of psychological, pharmacological, or combination interventions to treat anxiety compared with control conditions (eg, placebo, usual care, wait list or attention control conditions) among primary care patients were included. Intervention trials that recruited participants with either anxiety or depression among primary care patients were also included. Existing systematic reviews of psychological, pharmacological, or combination interventions were included for estimates of effect for general populations (ie, not limited to primary care populations). A decision tool developed by Pollock et al⁹ was adapted to identify the most current and comprehensive evidence.

Data Extraction and Quality Assessment

Two independent investigators rated the quality of studies as “good,” “fair,” or “poor,” using predefined criteria for each study type, in accordance with the USPSTF methods⁶ (eTable 1 in the Supplement). Discrepancies between raters were resolved by discussion or by consultation with the larger review team. Studies rated as “poor” quality due to critical methodological limitations were excluded, to limit the risk of bias in the included evidence.

Data from each included study were extracted into detailed forms using DistillerSR (Evidence Partners). One reviewer completed primary data abstraction, and a second reviewer checked all data for accuracy and completeness. Study inclusion criteria, population characteristics, intervention or screening tool details, comparators or reference standard details, and results for a priori-defined outcomes were extracted.

Data Synthesis and Analysis

Findings were synthesized using text, tables, and figures. Where possible, quantitative syntheses of test accuracy and anxiety treatment studies were conducted with meta-analysis. For meta-analysis of anxiety treatment (KQ4), the restricted maximum likelihood model with the Knapp-Hartung correction for small
numbers of studies was used.\textsuperscript{10,11} When studies included multiple intervention groups, the single most intensive or comprehensive intervention group per study was included in the meta-analysis. Standardized mean difference between groups (Hedges g) was the measure used for analysis, based on change from baseline in each group. Cohen rules of thumb were used to characterize standardized effect sizes of 0.20 as small, 0.50 as medium, and 0.80 as large.\textsuperscript{12}

In addition to presenting overall results, analyses were stratified by the presence of anxiety as an inclusion requirement. Studies in which all participants were required to meet some criteria for anxiety were shown separately from studies in mixed populations of people with anxiety or depression.

For meta-analysis of KQ2, data from 2 × 2 contingency tables were analyzed using a bivariate model, which modeled sensitivity and specificity simultaneously if possible. If there were not enough studies to use the bivariate model, sensitivity and specificity were pooled separately, using random-effects models with the method of DerSimonian and Laird.\textsuperscript{13} Point estimates were deemphasized when pooling fewer than 3 studies. For all analyses, statistical heterogeneity was assessed using the $I^2$ statistic. Analyses were conducted in Stata 16.1 (StataCorp). Significance testing was 2-sided, and results were considered statistically significant if $P \leq .05$.

The overall strength of the evidence for each KQ was assessed as high, moderate, low, or insufficient based on the overall quality of the studies, consistency of results between studies, precision of findings, and limitations of the body of evidence, using methods developed for the USPSTF.\textsuperscript{6} Additionally, the applicability of the findings to US primary care populations and settings was assessed. Discrepancies in assessments between team members were resolved by discussion.

**Results**

Altogether, 59 publications were included: 40 original studies (N = 275,489) and 19 existing systematic reviews (including ~483 studies [N = 815,07]) (Figure 2).
Reasons for exclusion: Intervention: Study used an excluded intervention/screening approach. Population: Study was not conducted in an average-risk population. Setting: Study was not conducted in a country relevant to US practice. Aim: Study aim not relevant. Outcomes: Study did not have relevant outcomes or had incomplete outcomes. Comparator: Study included a comparator group that was not included. Quality: Study did not meet criteria for fair or good quality. Design: Study did not use an included design. Existing systematic review superseded: Existing systematic review was superseded by one that was more contemporary, comprehensive, or relevant.

Benefits of Screening

KQ1. Do anxiety screening programs in primary care or comparable settings result in improved health outcomes in adults, including pregnant and postpartum persons?  
KQ1a. Does sending anxiety screening test results to providers (with or without additional care management supports) result in improved health outcomes?

Two RCTs (reported in 4 publications) examined the benefits of screening for anxiety in general adult populations14-17 (eTable 2 in the Supplement). One of these used a 5-item screener that included a single item for each of 5 conditions: anxiety, depression, pain, sleep disturbance, and fatigue.15 This study provided clinicians with a graphical depiction of T-scores from a follow-up instrument (the Patient-Reported Outcomes Measurement Information System), showing symptom levels and highlighting symptoms crossing a threshold for clinical importance. The other study used anxiety-related items from the Symptom Checklist-90-Revised to screen for anxiety alone; other conditions were not evaluated.16 Physicians in that study were trained in interpretation of the SCL-90-R and in management of anxiety.

Both trials found no reduction in anxiety symptoms or general psychological symptom severity compared with usual care at 13 to 22 weeks’ follow-up. The study that screened for anxiety along with other conditions reported a difference between groups in improvement of only 0.83 (standard error not reported) points on a 16-point anxiety scale at 3 months’ follow-up (P = .47).15 That study also found almost identical absolute change in its primary outcome of the General Severity Index, a measure of mental health symptom severity (−3.8 [SD, 8.5] in the intervention group vs −3.7 [SD, 8.7] in the control group; between-group difference, −0.1 [95% CI, −2.3 to 2.1]; P = .74). The study that screened for anxiety alone found no differences between groups at follow-up in anxiety symptom levels or in any of the 36-item Short Form Health Survey subscale scores at 5 months’ follow-up.16
Accuracy of Screening

**KQ2.** Do instruments to screen for anxiety accurately identify adults, including pregnant and postpartum persons, with anxiety risk in primary care or comparable settings?

Ten primary studies (in 12 articles; eTable 3 in the Supplement) reported the test accuracy of screening for anxiety with the GAD scale, Geriatric Anxiety Scale, Edinburgh Perinatal Depression Scale anxiety subscale, or Patient Health Questionnaire–Panic Disorder to detect generalized anxiety disorder, panic disorder, social anxiety disorder, or any anxiety disorder relative to a structured or semi-structured diagnostic interview administered within 2 weeks of the screening test (Figure 3). The most commonly studied instruments were the GAD-2 (range, 0-6) and the GAD-7 (range, 0-21), which demonstrated adequate accuracy for detecting generalized anxiety disorder. For example, in 3 studies the GAD-7 had a pooled sensitivity of 0.79 (95% CI, 0.69 to 0.94) and a specificity of 0.89 (95% CI, 0.83 to 0.94) to detect generalized anxiety disorder at a cutoff of 10 or greater (eFigure in the Supplement). At a cutoff of 3 or greater (on a scale from 0-6), the GAD-2 accurately identified 69% to 86% of adults (including pregnant women) with generalized anxiety disorder and 83% to 91% without it (eTable 4 in the Supplement). The GAD-2 needed a lower cutoff to obtain similar test accuracy to detect any anxiety disorder, with a cutoff of 1 or greater accurately identifying a similar proportion of those with any anxiety disorder (70%-90%) but at the cost of lower accuracy for identifying those without any anxiety disorder (55%-64%) (eTable 5 in the Supplement). At a cutoff of 2 or greater, the GAD-2 accurately detected 50% to 91% of adults with a panic disorder and 63% to 74% of those without a panic disorder (eTable 6 in the Supplement). At the same cutoff, the GAD-2 identified 85% of adults with social anxiety disorder and 62% of those without as well as or better than the GAD-2.

Harms of Screening

**KQ3.** What are the harms associated with screening for anxiety risk in primary care or comparable settings in adults, including pregnant and postpartum persons?

Neither of the 2 studies included for benefit of anxiety screening reported on harms. There was no pattern of effects indicating that screening might paradoxically increase anxiety or mental health symptoms.15,96

Benefits of Treatment

**KQ4.** Does treatment of anxiety risk result in improved health outcomes in adults, including pregnant and postpartum persons?

Twenty-six trials (reported in 36 publications) among primary care patients (eTables 8-9 in the Supplement) and 18 existing systematic reviews (not limited to primary care populations) addressed treatment for anxiety. Twenty-four of the included RCTs (n = 5307) examined psychological interventions and 2 (n = 423) examined pharmacological interventions. Among studies of psychological interventions, 14 included mixed populations of people with anxiety or depression, and 10 were limited to people with anxiety. Psychological interventions showed a relatively small but statistically significant reduction in anxiety symptom severity in primary care patients with anxiety (standardized mean difference [SMD], −0.41 [95% CI, −0.58 to −0.23]; 10 RCTs [n = 2075]; P2 = 40.2%). However, the effect was smaller and not statistically significant among mixed populations of people with anxiety or depression (SMD, −0.18 [95% CI, −0.39 to 0.03]; 12 RCTs [n = 1868]; P2 = 66.7%) (Table 1, Figure 4). The overall pooled effect size for all 22 studies was statistically significant, in favor of the intervention groups (SMD, −0.29 [95% CI, −0.44 to −0.15]; 22 RCTs [n = 3943]; P2 = 70.6%).

Psychological treatment was associated with reduced anxiety symptoms among the existing systematic reviews, which included an estimated 144 RCTs and approximately 11 000 participants. For example, SMDs at posttreatment follow-up among general adult populations would be considered large, as they were −0.80 and larger (eg, among people with generalized anxiety disorder: SMD, −0.80 [95% CI, −0.93 to −0.67]; 31 RCTs; N and P2 not reported) (eTable 12 in the Supplement). Psychological treatment was also associated with improved depression symptom severity and quality of life (eTable 13 in the Supplement). More limited evidence suggested a benefit in older and perinatal patients as well (eTable 12 in the Supplement).

Only 2 RCTs of pharmacotherapy in primary care patients met criteria for inclusion. These trials evaluated venlafaxine44 and escitalopram,45 and both showed a benefit. In the trial of venlafaxine, participants taking venlafaxine showed greater improvement in the primary outcome of anxiety symptoms at 24 weeks' follow-up, compared with placebo (mean difference at follow-up, −2.1 [95% CI, −4.2 to 0]; P = .05) (eTable 14 in the Supplement). In the RCT of escitalopram, which was limited to older adults, more participants taking escitalopram met the criteria for a treatment response than those taking a placebo (odds ratio, 1.87 [95% CI, 1.03 to 3.39]; 60% taking escitalopram compared with 45% taking a placebo had a treatment response, P = .05) (eTable 14 in Supplement).

Existing systematic reviews of general populations of patients reported improved anxiety and other outcomes for people taking antidepressants and benzodiazepines compared with placebo. For example, among patients with generalized anxiety disorder, the SMD for change in anxiety symptom severity with selective serotonin reuptake inhibitors was −0.66 (95% CI, −0.90 to −0.43); 23 studies (n = 2142); P2 not reported (eTable 15 in the Supplement). For antidepressants, benefits were seen for a variety of anxiety outcomes among people with generalized anxiety disorder, social anxiety disorder, and panic disorder. Limited evidence suggested that antidepressants and benzodiazepines may improve anxiety symptoms in older adults, but evidence in perinatal patients was lacking. Improvements were also seen for depression and social functioning outcomes with pharmacotherapy.

Harms of Treatment

**KQ5.** What are the harms of treatment of anxiety risk (psychotherapy or pharmacotherapy) in adults, including pregnant and postpartum persons?

None of the RCTs or existing systematic reviews of psychological treatment reported on adverse events, but there was no pattern of effects indicating an elevated risk of harm. For the harms of pharmacologic treatment, 3 RCTs (eTable 9 in the Supplement) and 8 existing systematic reviews addressing...
### Figure 3. Summary of Test Accuracy of Screening Instruments to Detect Anxiety Disorders (Key Question 2)

<table>
<thead>
<tr>
<th>Condition, screening test</th>
<th>Population</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Cutoff</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generalized anxiety disorder</strong></td>
<td>Adults</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1307</td>
<td>≥3</td>
<td>0.76 (0.68-0.85)</td>
<td>0.88 (0.87-0.88)</td>
</tr>
<tr>
<td></td>
<td>Pregnant women</td>
<td>1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>9750</td>
<td>≥3</td>
<td>0.69 (0.64-0.73)</td>
<td>0.91 (0.90-0.91)</td>
</tr>
<tr>
<td></td>
<td>Adults</td>
<td>3</td>
<td>2272</td>
<td>≥10</td>
<td>0.79 (0.65-0.94)</td>
<td>0.89 (0.83-0.94)</td>
</tr>
<tr>
<td><strong>Any anxiety disorder</strong></td>
<td>Adults</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1307</td>
<td>≥2</td>
<td>0.74 (0.69-0.79)</td>
<td>0.74 (0.70-0.78)</td>
</tr>
<tr>
<td></td>
<td>Pregnant women</td>
<td>2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>10474</td>
<td>≥1</td>
<td>0.79 (0.60-0.99)</td>
<td>0.64 (0.63-0.65)</td>
</tr>
<tr>
<td></td>
<td>Adults</td>
<td>3</td>
<td>1357</td>
<td>≥6</td>
<td>0.64 (0.46-0.82)</td>
<td>0.82 (0.79-0.87)</td>
</tr>
<tr>
<td></td>
<td>Pregnant women</td>
<td>1</td>
<td>954</td>
<td>≥6</td>
<td>0.57 (0.39-0.75)</td>
<td>0.87 (0.84-0.89)</td>
</tr>
<tr>
<td><strong>Panic disorder</strong></td>
<td>Adults</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1115</td>
<td>≥2</td>
<td>0.73 (0.34-1.0)</td>
<td>0.68 (0.57-0.79)</td>
</tr>
<tr>
<td></td>
<td>Adults</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1115</td>
<td>≥6</td>
<td>0.85 (0.71-0.98)</td>
<td>0.71 (0.56-0.86)</td>
</tr>
<tr>
<td></td>
<td>PHQ-PD</td>
<td>1</td>
<td>585</td>
<td>5</td>
<td>0.81 (0.69-0.93)</td>
<td>0.99 (0.98-1.0)</td>
</tr>
<tr>
<td><strong>Social anxiety disorder</strong></td>
<td>Adults</td>
<td>1</td>
<td>965</td>
<td>≥2</td>
<td>0.85 (0.73-0.93)</td>
<td>0.62 (0.59-0.65)</td>
</tr>
<tr>
<td></td>
<td>Adults</td>
<td>1</td>
<td>965</td>
<td>≥6</td>
<td>0.87 (0.75-0.94)</td>
<td>0.63 (0.60-0.66)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Pooled results for fewer than 3 studies shown only for illustrative purposes.

<sup>b</sup> For 1 study of pregnant patients examining the accuracy of the 2-item Generalized Anxiety Disorders scale (GAD), the total number of participants includes extrapolation to a larger sample (n = 9750) based on direct measurement of 528 participants (all those who screened positive and a random sample of those who screened negative). PHQ-PD indicates Patient Health Questionnaire–Panic Disorder.
Limitations due to small sample sizes (eTable 16 in the Supplement). Serious adverse events were rare, adverse events if they were taking medication (vs placebo) pants experiencing any adverse events and withdrawals due to nonserious harms as measured by a higher percentage of participants experiencing any adverse events and withdrawals due to adverse events if they were taking medication (vs placebo) (eTable 17 in the Supplement). Serious adverse events were rare, and data were insufficient to determine whether the risk of serious harms was increased. Case-control studies found an association between benzodiazepine use and suicide death85 and spontaneous abortion86 (eTable 18 in the Supplement). However, the studies’ inability to fully match cases and controls on severity of mental health symptoms and other health behaviors such as substance use limited confidence in the causal nature of these associations.

### Discussion

Evidence on the benefits and harms of screening programs for anxiety was limited and inconclusive. In contrast, a substantial evidence base indicated that effective treatments are available to treat anxiety, particularly cognitive behavioral therapy (CBT), antidepressants, and benzodiazepines (Table 2). The accuracy of the GAD-2 and GAD-7 was adequate to detect generalized anxiety disorder, but evidence on the test accuracy of screening tools had minimal replication for anxiety disorders other than generalized anxiety disorder. Because there are many disorders that manifest with anxiety symptoms (eg, posttraumatic stress disorder, obsessive-compulsive disorder, attention-deficit/hyperactivity disorder, depression, autism-spectrum disorders), sensitivity may be the more important than specificity when evaluating these tools. If screening tools identify patients with other conditions that need treatment in addition to anxiety disorders, there could still be a net value of screening, even at low specificity for anxiety alone.

### Anxiety Treatment

Evidence indicated that treatment for anxiety disorders is effective, including in populations with social anxiety disorder, panic disorder, or generalized anxiety disorder and in mixed populations with any of these anxiety disorders and depression. Effectiveness with these mixed populations is important to consider, since anxiety and depressive disorders often co-occur.85 Evidence also supported a benefit of psychological treatment among primary care patients, albeit with a smaller effect size than that for anxiety treatment overall. An independent review found a standardized mean difference of −0.39 (95% CI, −0.63 to −0.15) for primary care patients with depression or anxiety treated with CBT.88 This effect size is slightly larger than the findings of −0.29 (95% CI, −0.44 to −0.15) combining all studies (including those that included mixed populations with either depression or anxiety) and similar to the finding of −0.41 (95% CI, −0.58 to −0.23) when limited to individuals with anxiety. Differences in effect sizes between the 2 reviews may be partially explained by the fact that the independent review included some studies excluded from the current review because they were limited to people with certain medical conditions or because the studies received poor-quality ratings.

Most of the primary studies of anxiety interventions were conducted outside the US. Most participants included were White, and most studies targeted general adult (vs older adult or perinatal) populations. Most studies used CBT-based interventions, and few studies directly involved primary care clinicians in the delivery of treatment.

Potential pharmacological treatments for anxiety include antidepressants (particularly selective serotonin reuptake inhibitors and serotonin and norepinephrine reuptake inhibitors), antihistamines (such as hydroxyzine), β-blockers (such as propranolol), and anticonvulsant medications (such as gabapentin).89 Benzodiazepines, such as alprazolam or clonazepam, are often prescribed for acute anxiety or panic attacks. Buspirone is often used as an alternative to benzodiazepines because it does not carry a risk of

---

Table 1. Summary of Meta-analysis Results for Anxiety Outcomes in Primary Research Studies of Psychological Treatment of Anxiety in Primary Care Patients (Key Question 4)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. studies (No. analyzed)</th>
<th>Pooled result, SMD (95% CI)*</th>
<th>$I^2$, %</th>
<th>$T^2$</th>
<th>Range of effects (in native units)b</th>
<th>Median (IQR) effectsb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety symptom severity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All studies</td>
<td>22 (1943)</td>
<td>−0.29 (−0.44 to −0.15)</td>
<td>70.6</td>
<td>0.06</td>
<td>~8.0 to 6.8</td>
<td>~1.8 (~2.8 to −0.5)</td>
</tr>
<tr>
<td>Anxiety required</td>
<td>10 (2075)</td>
<td>−0.41 (−0.58 to −0.23)</td>
<td>40.2</td>
<td>0.02</td>
<td>~8.0 to 6.8</td>
<td>~2.3 (~3.0 to −1.4)</td>
</tr>
<tr>
<td>Anxiety or depression</td>
<td>12 (1868)</td>
<td>−0.18 (−0.39 to 0.03)</td>
<td>66.7</td>
<td>0.06</td>
<td>~6.1 to 4.5</td>
<td>~0.7 (~2.4 to 0.4)</td>
</tr>
<tr>
<td>Depression symptom severity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All studies</td>
<td>22 (1970)</td>
<td>−0.32 (−0.46 to −0.19)</td>
<td>66.4</td>
<td>0.05</td>
<td>~9.0 to 6.3</td>
<td>−1.50 (~2.6 to 0.01)</td>
</tr>
<tr>
<td>Anxiety required</td>
<td>9 (1990)</td>
<td>−0.49 (−0.74 to −0.25)</td>
<td>68.4</td>
<td>0.05</td>
<td>~9.0 to 6.3</td>
<td>−2.0 (~2.7 to −1.5)</td>
</tr>
<tr>
<td>Anxiety or depression</td>
<td>13 (1980)</td>
<td>−0.20 (−0.34 to −0.06)</td>
<td>39.9</td>
<td>0.02</td>
<td>~6.5 to 4.4</td>
<td>~0.7 (~2.4 to 0.01)</td>
</tr>
<tr>
<td>Mental Components score</td>
<td>7 (2104)</td>
<td>0.17 (−0.03 to 0.36)</td>
<td>54.4</td>
<td>0.02</td>
<td>~5.4 to 9.8</td>
<td>0.4 (~1.3 to 3.5)</td>
</tr>
<tr>
<td>Physical Component score</td>
<td>5 (1656)</td>
<td>0.03 (−0.12 to 0.18)</td>
<td>13.0</td>
<td>0.0</td>
<td>~1.5 to 2.2</td>
<td>0.3 (~1.5 to 0.6)</td>
</tr>
</tbody>
</table>

Abbreviation: SMD, standardized mean difference.

* Effect based on restricted maximum likelihood model with the Knapp-Hartung adjustment for small samples.

b Range of effects for all study groups, subgroup analyses, and time points, ie, not limited to records in the meta-analysis.
**Figure 4. Difference Between Groups in Change From Baseline in Anxiety Symptoms, for Primary Studies of Psychological Intervention for Treatment of Anxiety in Primary Care Populations Reported in Primary Randomized Clinical Trials (Key Question 4)**

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Intervention</th>
<th>Anxiety Outcome</th>
<th>No. randomized</th>
<th>Follow-up, wk</th>
<th>Mean (SD)</th>
<th>SMD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intervention</td>
<td>Control</td>
</tr>
<tr>
<td>Clark,62 2022</td>
<td>General</td>
<td>CBT</td>
<td>Any GAD-7</td>
<td>102</td>
<td>13</td>
<td>-5.3 (4.7)</td>
<td>-0.4 (5.5)</td>
</tr>
<tr>
<td>Nordgren,47 2014</td>
<td>General</td>
<td>CBT</td>
<td>Any BAI</td>
<td>100</td>
<td>10</td>
<td>-9.4 (8.9)</td>
<td>-5 (8.9)</td>
</tr>
<tr>
<td>Linden,46 2005</td>
<td>General</td>
<td>CBT</td>
<td>GAD HAM-A</td>
<td>72</td>
<td>14.5</td>
<td>-9.5 (9.6)</td>
<td>-1.5 (8.6)</td>
</tr>
<tr>
<td>Roy-Byrne,31 2010</td>
<td>General</td>
<td>CBT, medication, or both</td>
<td>Any BSI-12</td>
<td>1004</td>
<td>26</td>
<td>-7.2 (8.5)</td>
<td>-4.6 (8.9)</td>
</tr>
<tr>
<td>Fletcher,14 2005</td>
<td>General</td>
<td>CBT</td>
<td>Any HADS-A</td>
<td>30</td>
<td>12</td>
<td>-1.8 (2.8)</td>
<td>-1.8 (2.9)</td>
</tr>
<tr>
<td>Gemsichen,40 2019</td>
<td>General</td>
<td>CBT</td>
<td>PD BAI</td>
<td>419</td>
<td>26</td>
<td>-8.5 (13.2)</td>
<td>-5.3 (13.9)</td>
</tr>
<tr>
<td>Vera,61 2021</td>
<td>General</td>
<td>CBT</td>
<td>GAD DASS-21 anxiety</td>
<td>60</td>
<td>28</td>
<td>-12.5 (12.1)</td>
<td>-4.7 (10.2)</td>
</tr>
<tr>
<td>Stanley,35 2009</td>
<td>Older</td>
<td>CBT</td>
<td>GAD GADSS</td>
<td>134</td>
<td>26</td>
<td>-2.8 (3.8)</td>
<td>-1.6 (4.2)</td>
</tr>
<tr>
<td>Stanley,34 2014</td>
<td>Older</td>
<td>CBT</td>
<td>GAD GADSS</td>
<td>223</td>
<td>26</td>
<td>-2.9 (4)</td>
<td>-0.7 (4.5)</td>
</tr>
<tr>
<td>O’Mohen,63 2022</td>
<td>Perinatal</td>
<td>CBT</td>
<td>Any GAD-7</td>
<td>114</td>
<td>34</td>
<td>-5.1 (4.3)</td>
<td>-3.8 (4.7)</td>
</tr>
</tbody>
</table>

**Anxiety or depression**

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Intervention</th>
<th>Anxiety Outcome</th>
<th>No. randomized</th>
<th>Follow-up, wk</th>
<th>Mean (SD)</th>
<th>SMD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intervention</td>
<td>Control</td>
</tr>
<tr>
<td>Corsas,2 2022</td>
<td>General</td>
<td>CBT</td>
<td>Any GAD-7</td>
<td>105</td>
<td>8</td>
<td>-3.7 (4.4)</td>
<td>-1.4 (4)</td>
</tr>
<tr>
<td>Rollman,30 2018</td>
<td>General</td>
<td>CBT</td>
<td>Any PROMIS-anxiety</td>
<td>704</td>
<td>26</td>
<td>-9 (15)</td>
<td>-6.6 (8.7)</td>
</tr>
<tr>
<td>Lang,53 2006</td>
<td>General</td>
<td>PST</td>
<td>Any BSI-A</td>
<td>62</td>
<td>30</td>
<td>-5.4 (10.1)</td>
<td>-1.5 (8.7)</td>
</tr>
<tr>
<td>Graham,33 2020</td>
<td>General</td>
<td>CBT</td>
<td>Any GAD-7</td>
<td>146</td>
<td>8</td>
<td>-4.8 (4.9)</td>
<td>-1.4 (4.3)</td>
</tr>
<tr>
<td>Schreuders,32 2007</td>
<td>General</td>
<td>CBT</td>
<td>Any HADS-A</td>
<td>175</td>
<td>13</td>
<td>-1.5 (3.7)</td>
<td>-1.4 (3.4)</td>
</tr>
<tr>
<td>Proudfoot,44 2004</td>
<td>General</td>
<td>CBT</td>
<td>Any BAI</td>
<td>274</td>
<td>21</td>
<td>-8.7 (9.7)</td>
<td>-9 (8.7)</td>
</tr>
<tr>
<td>Sundquist,39 2015</td>
<td>General</td>
<td>MBT</td>
<td>Any SCL-ASS8</td>
<td>215</td>
<td>8</td>
<td>-0.5 (4)</td>
<td>-0.5 (7)</td>
</tr>
<tr>
<td>Seekles,33 2011</td>
<td>General</td>
<td>PST, case management</td>
<td>Any HADS-A</td>
<td>120</td>
<td>8</td>
<td>-1 (2.7)</td>
<td>-0.5 (2.9)</td>
</tr>
<tr>
<td>Lam,54 2010</td>
<td>Older</td>
<td>PST</td>
<td>Any HADS-A</td>
<td>299</td>
<td>26</td>
<td>-1.3 (3.8)</td>
<td>-1.7 (3)</td>
</tr>
<tr>
<td>Torres-Platas,39 2019</td>
<td>Older</td>
<td>MBT</td>
<td>Any GAD-7</td>
<td>61</td>
<td>8</td>
<td>-6.4 (5)</td>
<td>-2.3 (8)</td>
</tr>
<tr>
<td>Burger,34 2020</td>
<td>Perinatal</td>
<td>CBT</td>
<td>Any STAI</td>
<td>282</td>
<td>26 (postpartum)</td>
<td>-6.5 (11.9)</td>
<td>-7.7 (10.4)</td>
</tr>
<tr>
<td>Suchan,35 2002</td>
<td>Perinatal</td>
<td>CBT</td>
<td>Any GAD-7</td>
<td>63</td>
<td>13</td>
<td>-7.6 (4.6)</td>
<td>-3.4 (5.4)</td>
</tr>
</tbody>
</table>

**Anxiety or depression**

- Heterogeneity: r² = 0.06; I² = 66.66%; H² = 3.00
- Test of Θi = Θj: Q(9) = 29.64, P < .001
- Overall: -0.18 (-0.39 to 0.00)

- Heterogeneity: r² = 0.06; I² = 70.62%; H² = 3.40
- Test of group differences: Q(1) = 4.14, P = .04

The size of the data markers indicates the weight of each study in the analysis. BAI indicates Beck Anxiety Inventory; BSI, Brief Symptom Inventory; BSI-A, Brief Symptom Inventory-Anxiety; CBT, cognitive behavioral therapy; DASS, Depression Anxiety Stress Scales; GAD, Generalized Anxiety Disorder scale; GADSS, Generalized Anxiety Disorder Severity Scale; HADS-A, Hospital Anxiety and Depression Scale-Anxiety; HAM-A, Hamilton Anxiety Rating Scale-Anxiety; MBT, mindfulness-based therapy; PD, panic disorder; PROMIS-Anxiety, Patient-Reported Outcomes Measurement Information System-Anxiety; PST, problem solving therapy; SCL-ASS8, Symptom Checklist-Anxiety Symptom Scale; SMD, standardized mean difference; and STAI, State Trait Anxiety Inventory.
Table 2. Summary of Evidence: Anxiety Screening

<table>
<thead>
<tr>
<th>No. of studies</th>
<th>Summary of findings</th>
<th>Consistency and precision</th>
<th>Other limitations</th>
<th>Strength of evidence</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KQ1: Screening benefits</strong></td>
<td>2 RCTs (n = 918) Both studies found no group differences in anxiety or general mental health symptom severity at 13 to 22 wk of follow-up Absolute differences in change ranged from −1.5 to 0.3 on 16- and 40-point scales</td>
<td>Reasonably consistent, imprecise</td>
<td>Limited No. of studies</td>
<td>Insufficient</td>
<td>Both conducted in US primary care settings; 1 study published in 1994, so may not reflect current practice</td>
</tr>
<tr>
<td><strong>KQ2: Accuracy of screening tools</strong></td>
<td>10 Test accuracy studies (n = 6463) Adequate sensitivity and specificity for the GAD-7 to detect generalized anxiety disorder More limited evidence for the GAD-2 to detect generalized anxiety disorder GAD-7 and GAD-2 were less accurate for identifying any anxiety disorder Limited evidence for the GAD-7, GAD-2, and PHQ-PD to detect panic disorder Limited evidence for the GAD-7 and GAD-2 to detect social anxiety disorder</td>
<td>Reasonably consistent, reasonably precise</td>
<td>Few studies, limited replication</td>
<td>Moderate for the GAD-2/GAD-7 to detect generalized anxiety disorder Low for all other instruments and conditions</td>
<td>Many studies were conducted in the US, but those limited to older adults and pregnant women and the largest general adult study were conducted outside of the US</td>
</tr>
<tr>
<td><strong>KQ3: Harms of screening</strong></td>
<td>No studies reported on harms of screening for anxiety Studies included for KQ1 did not show a pattern of results indicating harmful impact</td>
<td>Consistent, imprecise</td>
<td>Minimal evidence</td>
<td>Insufficient</td>
<td>Both studies included for KQ1 outcomes conducted in US primary care settings; 1 study published in 1994 so may not reflect current practice</td>
</tr>
<tr>
<td><strong>KQ4: Benefits of treatment</strong></td>
<td>Psychological: 24 RCTs (n = 5307); 8 existing systematic reviews (n = 11 030) Psychological interventions showed a relatively small but statistically significant reduction in anxiety symptom severity in primary care patients with anxiety (SMD, −0.41 [95% CI, −0.58 to −0.23]; 10 RCTs [n = 2075]; I² = 40.2%) but not among mixed populations of people with anxiety or depression (SMD, −0.18 [95% CI, −0.39 to 0.03]; 12 RCTs [n = 1868]; I² = 66.7%)</td>
<td>Consistent, reasonably precise</td>
<td>Only 10 studies were among patients with anxiety, others were in mixed populations with anxiety or depression; limited evidence in older adults, limited evidence in perinatal patients; little information on outcomes beyond 8-12 wk</td>
<td>High for benefit</td>
<td>24 Studies in primary care populations, but only 7 conducted in the US All studies reporting race or ethnicity included majority (57% to 82%) White participants</td>
</tr>
<tr>
<td>Pharmacologic: 2 RCTs (n = 423); 10 existing systematic reviews (n = 227 RCTs [n = 40 803])</td>
<td>In the existing systematic reviews (not limited to primary care patients), psychological treatment was associated with reduced anxiety symptoms; SMDs at post treatment follow-up among broad adult populations were −0.80 and larger, and CBT was also associated with improved depression symptom severity and quality of life More limited evidence suggested a benefit in older and perinatal patients as well For pharmacologic treatment, 2 RCTs of venlafaxine and escitalopram in primary care patients both showed a benefit with antidepressant use Existing systematic reviews, not limited to primary care patients, reported improved anxiety and other outcomes for people taking antidepressants and benzodiazepines compared with placebo For example, among patients with generalized anxiety disorder, the SMD for change in anxiety symptom severity with SSRIs was −0.66 (95% CI, −0.90 to −0.43); 31 studies; N and I² not reported</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Population and settings characteristics were not reported in the existing systematic reviews. Low for psychological harms of pharmacotherapy, moderate for nonserious harms of pharmacotherapy, and insufficient for serious harms of pharmacotherapy.

Specific serious outcomes were rare, and studies were underpowered to identify; little information on outcomes beyond 8-12 weeks.

Harms Associated With Treatment for Anxiety

Antidepressants are widely used for the treatment of anxiety, and many people with anxiety have co-occurring depression disorders. Many of the existing systematic reviews included in the full report,7 which also covered depression and suicide risk screening, examined the risk of harm for any indication (including anxiety). Thus, many of the findings on antidepressant use for depression also apply to antidepressant use for anxiety, including a very small absolute increase in the risk of suicide and serious adverse events. Beyond antidepressants, very limited evidence on risk of serious harm with pharmacologic treatment for anxiety was identified, in both primary studies and existing systematic reviews. One included study found an association between use of benzodiazepines for treatment of anxiety and higher risk of suicide, but this was a relatively small case-control study that included information on 154 suicide deaths.85

Evidence on the risk of addiction or misuse of benzodiazepines was not reported in any studies included in the current report. However, the current evidence review identified a systematic review that examined studies reporting the association between benzodiazepines and suicide, although it did not meet quality criteria for inclusion in this review because it searched only 1 database and did not examine risk of bias (which is particularly important when synthesizing observational studies).90 However, it did identify 17 studies, most of which found an association between benzodiazepine use and suicide, covering a range of study populations.

In 2020, the US Food and Drug Administration (FDA) issued a warning:

...even when taken at recommended dosages, [benzodiazepine] use can lead to misuse, abuse, and addiction. Abuse and misuse can result in overdose or death, especially when benzodiazepines are combined with other medicines, such as opioid pain relievers, alcohol, or illicit drugs. Physical dependence can occur when benzodiazepines are taken steadily for several days to weeks, even as prescribed. Stopping them abruptly or reducing the dosage too quickly can result in withdrawal reactions, including seizures, which can be life-threatening.91

In addition, the FDA has issued a warning on the dangers of combined use of benzodiazepines with opioid medicines (including prescription pain and cough medications that contain opioids) and other central nervous system depressants.92 This combination can result in slowed or difficult breathing and death. Polypharmacy is of particular concern for older adults, who are more dependent. Despite the variety of treatment options, only 2 RCTs of pharmacotherapy in primary care patient populations were found. Both studies reported benefits of treatment with antidepressants (specifically, venlafaxine and escitalopram) for up to 24 weeks. Broad evidence from existing systematic reviews (not limited to primary care populations) also suggested improvements in anxiety and other outcomes (such as depression and social functioning) for a general adult population or older adults taking antidepressants or benzodiazepines for 1 to 3 months. Additional research is needed to address the benefit of pharmacological treatment for anxiety in perinatal populations.
likely to have multiple chronic conditions.\textsuperscript{93} While the absolute number of overdose deaths associated with prescription benzodiazepine use is low, it increased by 21% between 2019 and 2020, from 921 to 1122 per 100,000; 92.7% of these overdoses also involved opioids.\textsuperscript{94} Thus, while benzodiazepines are effective, based on the evidence included in this review, multiple streams of evidence suggested a need for caution and close monitoring of their use.

### Limitations

This review had several limitations. First, it was designed to establish whether there are effective treatments and valid screening tools feasible for use in primary care. Its scope did not include determining the accuracy of all possible anxiety screening instruments or the effect sizes of all specific types of psychological and pharmacologic treatments and their comparative effectiveness. Second, studies were excluded if they were conducted in narrow populations that were not widely applicable to screening in primary care settings but that are seen regularly in primary care settings, nevertheless. For example, studies were not included if they were limited to persons with physical or developmental disabilities or to people with medical or other mental health comorbidities such as heart disease, cancer, substance use disorders, bipolar disorder, or posttraumatic stress disorder.

Third, similarly, the screening instruments selected for review may not apply to some important groups of patients, such as those with low literacy, low health literacy, limited verbal language, or patients who do not speak English. Fourth, the review was limited to studies conducted in highly developed countries and has limited generalizability to low- and middle-income countries.

Fifth, only 2 studies were found of anxiety screening programs, one of which was published nearly 30 years ago; more studies comparing primary care-based anxiety screening with usual care are needed, particularly those using screening tools with evidence of diagnostic accuracy. Sixth, evidence on the accuracy of anxiety screening tools other than the GAD-2 or GAD-7 was also limited. Because this review focused on a limited number of screening tools, additional diagnostic accuracy studies may be available that were not included. However, informal searching indicated that evidence on other screening tools is very unlikely to provide more robust evidence for any instrument than was found for the GAD-2 and GAD-7 in this review.

### Conclusions

Evidence was insufficient to draw conclusions about the benefits or harms of anxiety screening programs. However, clear evidence exists that treatment for anxiety is beneficial, and more limited evidence indicates that some anxiety screening instruments have acceptable accuracy to detect generalized anxiety disorder.

### References

8. O’Connor E, Perdue LA, Coppola EL, Henninger ML, Thomas RG, Gaynes BN. Depression and...


77. Lader M, Scotto JC. A multicentre double-blind comparison of hydroxyzine, buspirone and placebo in patients with generalized anxiety disorder.


or
