Office-Based Screening of Common Psychiatric Conditions

Sirisha Narayana, MD\(^a\), Christopher J. Wong, MD\(^b\)*

**KEYWORDS**
- Screening • Depression • Anxiety • Cost-effectiveness • Outcomes

**KEY POINTS**
- Depression and anxiety disorders are common and significant conditions in the general population.
- Multiple well-validated screening instruments exist, which may be easily administered in an outpatient setting. These include the Patient Health Questionnaire (PHQ)-9 for depression, the Generalized Anxiety Disorder (GAD)-7 for anxiety disorders, and the Primary Care–Posttraumatic Stress Disorder Screen (PC-PTSD) for PTSD.
- Despite the availability of screening tools, the overall cost-effectiveness of general screening for anxiety or depression is uncertain.
- Screening for depression is recommended by some preventive health guidelines, and is most likely cost-effective in the setting of high prevalence and the availability of treatment using a collaborative care model.

**INTRODUCTION**
Depression and anxiety disorders are common and significantly affect health worldwide. Treatment options including psychotherapy and pharmacotherapy have expanded and in many regions are easily accessible. Yet these disorders may be undertreated. Approximately 40% of patients screening positive for anxiety disorders were not receiving treatment in one study, and patients with depression were being treated only 50% of the time in another study, with disparities among ethnic/racial groups.\(^1,2\)
Screening is therefore an important element to consider in the effort to reduce the overall burden of depression and anxiety disorders. Multiple screening modalities have been

Funding sources: None.
Conflicts of interest: None.
\(^a\) Division of General Internal Medicine, Department of Medicine, VA Puget Sound Health Care System, University of Washington, 1660 South Columbian Way, Seattle, WA 98108, USA;
\(^b\) Division of General Internal Medicine, Department of Medicine, University of Washington, 4245 Roosevelt Way Northeast, Box 354760, Seattle, WA 98105, USA
* Corresponding author.
E-mail address: cjwong@uw.edu

Med Clin N Am 98 (2014) 959–980
http://dx.doi.org/10.1016/j.mcna.2014.06.002
© 2014 Elsevier Inc. All rights reserved.
developed to facilitate diagnosis and treatment of common mental health disorders in the primary care setting. With the many options available, it is important to have an understanding of the strengths and limitations of these tools, the recommendations from major guidelines regarding screening, and where unanswered questions remain.

**SCREENING ASYMPTOMATIC PATIENTS FOR PSYCHIATRIC CONDITIONS—GENERAL CONSIDERATIONS**

Screening requires several conditions be present to be considered effective (Table 1). First, the illness should be significantly burdensome in the population to warrant screening. The reported prevalence of depression and anxiety disorders is high, although estimates vary by location, classification, and duration (Table 2 shows selected studies). Prevalence estimates should be interpreted with caution. Variations exist by country, and because patients with psychiatric disorders may incur more physician visits, clinic-based point prevalence estimates are generally higher than those using population-based methods (eg, generalized anxiety disorder had a 3.1% prevalence in a community sample vs 7.6% in a clinic-based sample). Second, a highly sensitive and specific screening test that is easy to administer must exist. Third, the illness should be identified by screening at a treatable stage or a stage in which early treatment is more effective than later treatment. The concept of early treatment is more complex with psychiatric illnesses: by definition patients are symptomatic, but the natural history of common psychiatric conditions is varied; they may have potentially lifelong

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Conceptual framework for psychiatric disease screening</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Criteria</strong></td>
<td><strong>Nonpsychiatry Examples</strong></td>
</tr>
<tr>
<td>Condition causes significant burden in the population</td>
<td>Rare but severe: phenylketonuria in newborns</td>
</tr>
<tr>
<td>An easy-to-administer, effective screening test exists</td>
<td>Blood pressure</td>
</tr>
<tr>
<td>Early treatment is more effective than later treatment</td>
<td>Cancer screening: goal is to identify disease at an earlier stage at which treatment is more effective Diabetes: goal is to identify disease before it is symptomatic to initiate treatment and prevent complications</td>
</tr>
<tr>
<td>Benefits of screening tests and subsequent treatment outweigh potential harms, at acceptable cost</td>
<td>Mammography: Harms include radiation, follow-up imaging, biopsies, worry. Optimal target population and interval still debated</td>
</tr>
</tbody>
</table>
conditions such as generalized anxiety disorder, with waxing and waning severity, or episodic with a self-limited course as with major depressive disorder. Fourth, the screening tests and treatment must have clinically meaningful benefits that outweigh potential harms to a patient at an acceptable cost to society. For the purposes of this review, the third and fourth criteria are considered together as general effectiveness.

**Case Definitions**

Unless specified otherwise, this review discusses screening tools validated against structured diagnostic interviews. Most of the tools described later were tested against the Diagnostic and Statistical Manual of Mental Disorders (DSM)-III, the DSM-IV, or World Health Organization (WHO) definitions. The newer DMS-V diagnostic criteria for a major depressive episode and generalized anxiety disorder have only minor differences compared with those of DSM-IV, not expected to substantially affect the performance of these measures.6,7

**DEPRESSION**

**Burden of Disease**

The 12-month prevalence of major depressive disorder in the United States is estimated at 6.7%, and the lifetime prevalence of any mood disorder is approximately 20%.5,8 Depression is estimated as the fourth leading cause of disability adjusted life years worldwide.4

**Screening Tools**

**General population**

Many screening tools exist for depressive disorders (Table 3). Selected tools are discussed below.

**PHQ-9 and PHQ-2** The PHQ-9 evolved from the full Primary Care Evaluation of Mental Disorders (PRIME-MD) instrument developed in the early 1990s to diagnose depression, anxiety, somatoform disorder, and alcohol and eating disorders.9 The PRIME-MD necessitated significant investment of physician-patient time (>11 minutes) and led to the shorter PHQ,10 and subsequently the PHQ-9, the 9-item depression module from the full PHQ (Box 1). It is scored 0 to 27, with a score of 10 or more indicating a possible depressive disorder. Scores of 5, 10, 15, and 20 represented mild, moderate,
<table>
<thead>
<tr>
<th>Population</th>
<th>Tools</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>General population/primary care</td>
<td>PHQ-9</td>
<td>PHQ-9 and PHQ-2 freely available</td>
</tr>
<tr>
<td></td>
<td>PHQ-2</td>
<td>Screening recommended by USPSTF if treatment resources available</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td></td>
</tr>
<tr>
<td>Veterans</td>
<td>PHQ-9</td>
<td>The Veterans Affairs administration recommends yearly screening of veterans for depression using the PHQ-2(^66) with follow-up as needed with PHQ-9 and a full assessment; and those at higher risk of developing depression (patients with hepatitis C beginning interferon treatment or patients post-MI) be given the PHQ-9 when depression is suspected Use caution in screening patients older than 75 y because instruments may not perform as well as between the ages of 65 and 75 y(^67)</td>
</tr>
<tr>
<td></td>
<td>PHQ-2</td>
<td></td>
</tr>
<tr>
<td>Medical comorbidities</td>
<td>Various, including PHQ-9, PHQ-2, HADS, BDI, Montgomery and Asberg Depression Rating Scale, others</td>
<td>Increased rates of depression in patients with central nervous system disease, cardiovascular disease, and malignancy(^68-71) Up to 20% of patients with myocardial infarction meet DSM-IV criteria for major depression. The American Heart Association advocates for screening for depression after MI. Approximately half of post-MI depression resolves without treatment. Some evidence of treatment efficacy in patients poststroke, post-MI(^68-75) The PHQ-9 and HADS have been studied in patients with coronary disease and are comparable in effectiveness for screening(^76) The PHQ-9 has been used in multiple medical settings, including rheumatology, ophthalmology, and spinal cord injury(^77)</td>
</tr>
<tr>
<td>Elderly</td>
<td>PHQ-9</td>
<td>GDS score &gt;5: sensitivity 0.92, specificity 0.81 for major depression</td>
</tr>
<tr>
<td></td>
<td>PHQ-2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GDS</td>
<td></td>
</tr>
<tr>
<td>Addiction</td>
<td>PHQ-9, others</td>
<td>Consider screening in patients with substance use disorders</td>
</tr>
<tr>
<td>Psychiatric comorbidities</td>
<td>PHQ-9, others</td>
<td>Consider screening in patients with other psychiatric conditions, including anxiety disorders and PTSD</td>
</tr>
</tbody>
</table>

Abbreviations: GDS, Geriatric Depression Scale; HADS, Hospital Anxiety and Depression Scale; MI, myocardial infarction.
moderately severe, and severe depression, respectively. To make the diagnosis of major depression, at least one of the first two questions must score a two or greater; this includes anhedonia and feelings of low mood/depression. The PHQ-9 module aligns closely with the DSM-IV diagnosis of depression and accordingly includes a question of whether symptoms impair functioning. At a cutoff score of 10, the PHQ-9 was found to have 0.88 sensitivity and specificity, and a likelihood ratio of 7.1, in the population studied (mean age 30–40 years, mostly white women with few medical comorbidities).11 The validity of the PHQ-9 has been further demonstrated in a general population.12 Subsequent meta-analyses confirmed high sensitivity of the PHQ-9, although one study did demonstrate a lower specificity of 0.77, possibly due to a lower prevalence of depression in the population studied.13 Despite the heterogeneity of the studies in terms of settings (community, primary care, and hospital specialties), the properties of the PHQ-9 for major depression were consistent across this range.14 The PHQ-2 consists of only the first 2 items of the PHQ-9; a PHQ-2 score of 3 or more has a sensitivity of 0.83 and a specificity of 0.92 for major depression.15 The increased specificity of the PHQ-2 may be better for screening larger populations. However, a comparison of these 2 instruments has not yet been conducted. In addition, the PHQ-9 may also be used to monitor treatment response.

Box 1
Patient health questionnaire-9

<table>
<thead>
<tr>
<th>Over the Last 2 wk, How Often Have You Been Bothered by Any of the Following Problems?</th>
<th>Not At all</th>
<th>Several Days</th>
<th>More Than Half the Days</th>
<th>Nearly Every Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Little interest or pleasure in doing things</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2. Feeling down, depressed, or hopeless</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3. Trouble falling or staying asleep, or sleeping too much</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4. Feeling tired or having little energy</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5. Poor appetite or overeating</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6. Feeling bad about yourself — or that you are a failure or have let yourself or your family down</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7. Trouble concentrating on things, such as reading the newspaper or watching television</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8. Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9. Thoughts that you would be better off dead or of hurting yourself in some way</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

0 + ______ + ______ + ______ =Total Score: ______

If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all Somewhat difficult Very difficult Extremely difficult

Other scales The Beck Depression Inventory for Primary Care is a 7-item scale adapted from the 21-item Beck Depression Inventory (BDI). It has a 0.97 sensitivity and 0.99 specificity for a score of 4 or more but requires a license fee for use. The World Health Organization Five is a 5-item scale that was found to be slightly more sensitive but less specific than the brief PHQ in a study of 400 primary care patients. There are increasing numbers of studies of its use in different populations, and it is also freely available online in all languages. The Mental Health Inventory is a 5-item mental health tool that was used as a comparison for the validity of the PHQ, but it is not specific for depression. Single-question screening methods have a low sensitivity of 0.32 but a high specificity of 0.97 and thus cannot be relied on as an effective screening tool.

Comparisons of the various screening instruments have been conducted. In 38 studies involving 32,000 patients in primary care settings, depression instruments were found to be comparably effective (this included PHQ-9, BDI, and the Geriatric Depression Screen, among others) with a median likelihood ratio positive of 3.3 and were quick to use (administration times ranged from <2 to 6 minutes).

Effectiveness of screening for depression

Data are conflicting as to whether general population– or primary care–based screening leads to improved patient outcomes. A systematic review of randomized trials of screening questionnaires administered by research assistants did not show any change in physician diagnoses or interventions; these studies used the BDI, General Health Questionnaire, and the Zung Self-Rating Depression Scale. A more recent review for the US Preventive Services Task Force (USPSTF) found that there was only one fair-quality randomized controlled trial that directly assessed whether screening for depression among adults in primary care reduces morbidity or mortality and showed mixed results. Based on data from four different trials, screening was thought to be effective when ancillary staff was involved in depression care and extra efforts were made to enroll patients in specialty mental health treatment. A Cochrane review in 2005 concluded that screening and case-finding instruments, when administered without any additional care structure, had little impact on the overall recognition rates of depression, management of depression (or intervention with antidepressants), and outcomes from depression. No evidence was found to address the harms of screening specifically. Potential harms include stigma and psychological effects of false-positive results of diagnoses and unnecessary treatment with and exposure to side effects of antidepressant medications. Some studies, but not all, showed concern for a possible increase in upper gastrointestinal bleeding in older adults when taking selective serotonin reuptake inhibitors (SSRIs) and nonsteroidal antiinflammatory medications together. There is no definitive evidence that suicidal behavior increases with second-generation SSRIs, although there is evidence that there is increased risk under the age of 25 years and decreased risk over the age of 65 years.

Screening for depression was found to be cost-effective only in settings of high prevalence of depression and high treatment and remission rates. Costs were high for annual screening but were lower for screening every 5 years and only truly cost-effective for one-time screening. Pearls for depression screening are listed in Box 2.

ANXIETY DISORDERS

Burden of Disease

Anxiety disorders have a 12-month prevalence of 18% and a lifetime prevalence as high as 29% (see Table 2). Morbidity includes a high degree of interference with life activities; increased number of physician visits, especially if there are somatic symptoms; and decreased functional status.
Screening Tools

**Beck anxiety inventory**

Selected tools and considerations are listed in **Table 4**. The Beck Anxiety Inventory (BAI) is a 21-item, patient-completed questionnaire, developed to discriminate anxiety disorders from depressive disorders in an outpatient psychiatric clinic. Its questions primarily report somatic symptoms. An abbreviated version, the Beck Anxiety Inventory-Primary Care (BAI-PC), has subsequently been developed, and, although...

---

### Table 4

**Selected tools and considerations for screening for anxiety disorders**

<table>
<thead>
<tr>
<th>Population</th>
<th>Tools</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>General population/primary care</td>
<td>GAD-7, GAD-2</td>
<td>GAD-7 and GAD-2 freely available</td>
</tr>
<tr>
<td>Medical comorbidities</td>
<td>GAD-7, others</td>
<td>Consider assessment for chronic somatic symptoms such as headache syndromes, chronic pain, gastrointestinal disorders</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Somatic symptoms such as gastrointestinal symptoms have been found to have a high prevalence of anxiety disorders in primary care</td>
</tr>
<tr>
<td>Addiction</td>
<td>GAD-7, GAD-2</td>
<td>Perform better as general screens for anxiety disorders than for GAD specifically</td>
</tr>
<tr>
<td>Elderly</td>
<td>GAD-7, Geriatric Anxiety Inventory Short Form (GAI-SF)</td>
<td>Consider using a lower cutoff score of 5–7 if using the GAD-7 GAI-SF: sensitivity 0.75, specificity 0.84</td>
</tr>
<tr>
<td>Psychiatric comorbidities</td>
<td>GAD-7, others</td>
<td>Consider screening for those with depressive disorders, other anxiety spectrum disorders, PTSD</td>
</tr>
</tbody>
</table>
a follow-up study has been completed, it has not been extensively retested using diagnostic interviews as the gold standard.\textsuperscript{27,28}

**Hospital anxiety and depression scale**
Hospital anxiety and depression scale (HADS) was developed to screen medical patients for psychiatric conditions. Despite its name, it has also been validated in primary care populations. The optimal cutoff score for the anxiety subscale (HADS-A) is approximately 8, with a sensitivity and specificity in the 0.70 to 0.90 range.\textsuperscript{29}

**Generalized anxiety disorder-7**
GAD-7 is another patient-completed questionnaire (Box 3).\textsuperscript{30} Unlike the BAI, its criteria closely mirror the DSM-IV definition of generalized anxiety disorder, with the exception that it asks for a symptom report for the prior 2 weeks rather than for 6 months. As would be expected, using higher cutoffs of the GAD-7 yields lower sensitivity but higher specificity. A follow-up study found that it also performed well in identifying PTSD, panic disorder, and social anxiety disorder, with a sensitivity of 0.80 and a specificity of 0.76 at a cutoff of 7; for GAD alone, a cutoff score of 10 maintained sensitivity while improving specificity.\textsuperscript{1} This tool has been shortened further to the GAD-2, a 2-item questionnaire for which a score of 3 has a sensitivity of 0.86 and a specificity of 0.83 for generalized anxiety disorder.\textsuperscript{1}

**Multistage screening tools**
The Symptom Driven Diagnostic System–Primary Care (SDDS-PC) was developed as a 16-item, patient-completed screening tool for multiple mental health disorders in primary care.\textsuperscript{31} In the initial study, it had 0.85 sensitivity and 0.60 specificity for generalized anxiety disorder, with lower sensitivity and higher specificity for panic disorder.\textsuperscript{31} Although the performance of SDDS-PC is comparable with that of GAD-7 as an initial screen, it was designed to necessitate nurse or physician follow-up using a proprietary assessment module. Similarly, the PRIME-MD tool, as discussed earlier for

---

**Box 3**

**Generalized anxiety disorder-7**

<table>
<thead>
<tr>
<th>Over the Last 2 wk, How Often Have You Been Bothered by the Following Problems?</th>
<th>Not At All</th>
<th>Several Days</th>
<th>More Than Half the Days</th>
<th>Nearly Every Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Feeling nervous, anxious or on edge</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2. Not being able to stop or control worrying</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3. Worrying too much about different things</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4. Trouble relaxing</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5. Being so restless that it is hard to sit still</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6. Becoming easily annoyed or irritable</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7. Feeling afraid as if something awful might happen</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

\[0 + \_ + \_ + \_ = \text{Total Score: }\]

If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all Somewhat difficult Very difficult Extremely difficult

depression, also contains a screen for anxiety disorders but was designed to be used with additional modules. The PRIME-MD later evolved into the PHQ and then the shorter PHQ-9 and GAD-7. Neither SDDS-PC nor PRIME-MD is freely available for routine use, and the efficacy of using only the screening portion of such tools for anxiety disorders in a real-world setting is unknown.

**Effectiveness of Screening for Anxiety**

There are few cost-benefit studies available to guide implementation of screening of anxiety and depression. Small ambulatory studies showed that screening for anxiety is feasible and led to increased diagnoses even in training clinics but lacked cost-effectiveness data. The collaborative care model is likely an effective intervention for both depression and anxiety, with data supporting cost-effectiveness for depression, but lacking adequate studies for other anxiety disorders. Cost-effectiveness data are needed, as screening may identify illness of lesser severity, whereas more severe illnesses may present clinically without the need for screening. The shorter screening tests such as the PHQ-9 and GAD-7 only take minutes for the patient to complete and the provider to review, and its perceived utility is acceptable to patients. Pearls for anxiety disorder screening are shown in Box 4.

**OTHER COMMON PSYCHIATRIC CONDITIONS**

Common mental health conditions encountered in the primary care office setting, including the neurodevelopmental disorder adult attention-deficit and hyperactivity disorder (ADHD) and the trauma-related disorder PTSD, are briefly discussed in the following sections. Substance use disorder (formerly substance use and dependence in DMS-IV) is not covered in this review.

**ADHD**

**Burden of Disease**

ADHD was originally thought to be primarily a pediatric disorder, whereas 40% to 60% of children with ADHD have symptoms that persist into adulthood. ADHD is thought to have a prevalence of approximately 4% in the adult population (see Table 2). In adulthood, symptoms of hyperactivity are less pronounced compared with childhood.

---

**Box 4**

**Screening for anxiety disorders: pearls**

1. Anxiety disorders are widely prevalent, and there are freely available screening tools for use in primary care settings.
2. Despite wide prevalence and availability of proven screening tools, there is no conclusive evidence to support cost-effectiveness of screening general populations for anxiety disorders.
3. The GAD-7, using a cutoff of 10 or more, is a reasonable screen for generalized anxiety disorder (sensitivity 0.89, specificity 0.82). If a shorter screen is desired, the 2-item GAD-2 at a cutoff of 3 or more is an option (sensitivity 0.86, specificity 0.83).
4. The GAD-7 also screens for other disorders, including panic disorder and PTSD; it is critical to follow up this screening tool with an accurate clinical assessment.
5. Although not specifically screening, consider assessment for anxiety disorders in patients presenting with unexplained somatic symptoms.
6. Consider screening in patients with other psychiatric conditions, including depression and PTSD, as well as substance use disorders.
with those of inattention. The DSM-V diagnosis of ADHD emphasizes pervasive symptoms of inattention, hyperactivity, and impulsivity affecting least 2 domains of daily life (eg, work and home). Diagnosis is based on clinical evaluation, which should include assessment of other psychiatric illnesses and substance abuse, impact of these symptoms on daily functioning, and a developmental history.

**Screening Tools**

Several self-reporting measures exist for ADHD, but none are sufficient for diagnosis alone. The New York University ADHD Program advocates the use of Adult Self-Report Scale V1.1 from WHO. It consists of a total of 18 questions that correlate to DSM diagnostic criteria for ADHD, and the screener portion has 6 questions that patients can self-report. In a small study of 154 respondents, the screener questions were found to have a sensitivity of 0.69 and a specificity of 1.0, with an accuracy of 98%. Subsequent validity studies showed strong concordance with clinician diagnoses.

The Wender Utah Rating Scale (WURS) was originally found to have a sensitivity of 0.86, but subsequent studies showed a sensitivity of 0.72 and a specificity of 0.58, suggesting that it misclassifies about half of those without ADHD. The Connors’ Adult ADHD Rating Scales (CAARS) has separate scales for the patient and for completion by an observer such as a spouse, friend, or parent, so that physicians can gather corroborative data. Each scale has a screening, short, and long form. The Current Symptoms Scale asks about adult patients’ behaviors in the last 6 months. It also has a separate scale for patients and for an observer. It is unclear whether any of these scales are superior. In a 2011 systematic review that included 14 scales for ADHD, CAARS and WURS had the best psychometric properties. Firm conclusions were limited because of the poor quality of many of the studies identified.

**Effectiveness of screening for ADHD**

More data are needed to clarify the value of screening for ADHD and the best tool for screening. Controversy exists about the diagnosis given that the symptoms of ADHD are challenging to differentiate from those of other psychiatric diagnoses or substance abuse. The Adult Self-Report Scale Screener shows considerable promise, and physicians should be alerted to its potential and for clinical signs of ADHD in their patients (Box 5).

**POSTTRAUMATIC STRESS DISORDER**

**Burden of Disease**

The overall lifetime prevalence of PTSD is approximately 7%, higher in veterans and other risk groups. PTSD is associated with functional impairment, increased mental health care utilization, and increased psychiatric comorbidities.

---

**Box 5**

**Screening for ADHD: pearls**

1. ADHD is now considered a neurodevelopmental disorder in the DSM-V, with a prevalence of approximately 4% in adults.

2. There are several screening tools published for ADHD for patient or caregiver/observer report. The Adult Self-Report Scale V1.1 from the World Health Organization is freely available.

3. Although there are available tools for screening for ADHD, there is not sufficient evidence with regard to burden of illness or cost-effectiveness to support screening general populations.
Screening Tools

Screening tools for conditions such as generalized anxiety disorder may also be effective at screening for PTSD. For the office-based setting, if screening for anxiety disorders is performed, then use of the GAD-7 is a reasonable option. There are numerous screening tools for PTSD specifically, with validation in different populations, including at-risk groups such as veterans. The US Department of Veterans Affairs endorses several screening tools, including the BAI-PC, PC-PTSD, Short Form of the PTSD Checklist–Civilian Version (PCL-C), Short Screening Scale for Short Post-Traumatic Stress Disorder Rating Interview (PTSD), Startle, Physiological arousal, Anger, and Numbness (SPAN), Short Post-Traumatic Stress Disorder Rating Interview (SPRINT), and the Trauma Screening Questionnaire. Of these, the 4-item PC-PTSD (Box 6) is available online through the US Department of Veterans Affairs Web site and has been widely used in the veterans affairs (VA) system, whereas the other tools are either proprietary or require request from the researchers.

The PC-PTSD has been validated in the primary care setting at the VA with a sensitivity of 0.78 and specificity of 0.87 at a recommended cutoff of 3 (out of 4). Alternatively, there are PTSD-specific screens that are available online that patients may complete on their own. The VA patient portal uses a version of the 17-item PCL specific for the veteran population. Each question has a 5-point Likert scale ranging from not at all to extremely for possible PTSD symptoms. The recommended cutoff ranges from 30 to 50, depending on the population studied and the

<table>
<thead>
<tr>
<th>Box 6</th>
<th>Primary care PTSD screen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>The PC-PTSD is a 4-item screen that was designed for use in primary care and other medical settings and is currently used to screen for PTSD in veterans at the VA. The screen includes an introductory sentence to cue respondents to traumatic events. The authors suggest that in most circumstances the results of the PC-PTSD should be considered “positive” if a patient answers “yes” to any 3 items. Those screening positive should then be assessed with a structured interview for PTSD. The screen does not include a list of potentially traumatic events.</td>
</tr>
<tr>
<td><strong>Scale</strong></td>
<td>Instructions: In your life, have you ever had any experience that was so frightening, horrible, or upsetting that, in the past month, you:</td>
</tr>
<tr>
<td></td>
<td>1. Have had nightmares about it or thought about it when you did not want to? YES/NO</td>
</tr>
<tr>
<td></td>
<td>2. Tried hard not to think about it or went out of your way to avoid situations that reminded you of it? YES/NO</td>
</tr>
<tr>
<td></td>
<td>3. Were constantly on guard, watchful, or easily startled? YES/NO</td>
</tr>
<tr>
<td></td>
<td>4. Felt numb or detached from others, activities, or your surroundings? YES/NO</td>
</tr>
<tr>
<td></td>
<td>Current research suggests that the results of the PC-PTSD should be considered “positive” if a patient answers “yes” to any three items.</td>
</tr>
</tbody>
</table>

gold standard used to evaluate it, although a repeat study found that a cutoff of 60 maximized diagnostic efficiency (percentage of cases correctly diagnosed), albeit at a low sensitivity of 0.56 and a specificity of 0.92. For screening purposes, if identifying the greatest percentage of cases is prioritized, a lower cutoff may be considered.

**Effectiveness of screening for PTSD**

Use of PTSD screens in general primary care populations outside the VA setting is less certain. Unlike the VA system, many primary care systems may not have access to specialized PTSD treatment; screening may not be desired if there is not readily available and effective treatment. These tools (PC-PTSD, PCL) may be considered if the clinical presentation includes symptoms suggestive of PTSD, although such use would not be strictly screening. Although some online Web sites use the 17-item questionnaire, the 4-item PC-PTSD may be a better initial screen for an in-office setting, especially if it would take additional time to complete in conjunction with screens for other conditions (Box 7).

**LIMITATIONS OF EXISTING EVIDENCE**

Despite the burden of disease and the availability of effective screening tools, there are still significant considerations in implementing these tools in clinical practice.

**Optimal Cutoffs for Screening Tools**

Most questionnaire-based screening tools use a quantitative scoring system. Operating characteristics such as the area under the curve vary depending on the population studied and the gold standard used and accordingly affect the optimal cutoff score. In addition, there is a value judgment in defining the optimal cutoff score. If one places a priority on sensitivity, assuming there is a readily available way to identify true diagnoses, then one would seek a sensitivity more than 0.9 and accept a lower specificity; the positive likelihood ratio of a positive result at such a cutoff would generally be low. Conversely, one may attempt to optimize both the sensitivity and specificity, accepting a lower sensitivity to preserve a higher specificity. Using the GAD-7 as an example, a cutoff score of 5 for screening of generalized anxiety disorder alone yields a sensitivity of 0.97 and a specificity of 0.57, whereas the suggested cutoff score of 10 still preserves reasonable sensitivity of 0.89 but improves the specificity markedly to 0.82. Thus, it is straightforward to recommend a cutoff score of 10; the small decrement in sensitivity, still at nearly 0.90, is an acceptable trade-off to improve the specificity to more than 0.80. However, if one uses the GAD-7 as a combined screen for generalized anxiety disorder, panic disorder, social anxiety disorder, and

---

**Box 7**

**Screening for PTSD: pearls**

1. PTSD has a significant burden and there are freely available screening tools.
2. Despite its prevalence, the best evidence for testing and most common strategies use a targeted screening approach.
3. Screening is likely most effective when performed where there are adequate resources for treatment of PTSD.
4. The PC-PTSD is a short, 4-item, freely available screening tool that has a sensitivity of 0.78 and specificity of 0.87 at a recommended cutoff of 3/4, best studied in veterans.
PTSD, the choice of cutoff is a bit more difficult; the cutoff score of 5 has a 0.90 sensitivity and 0.63 specificity, but increasing the cutoff score to 10 markedly diminished the sensitivity to 0.68 while increasing the specificity to 0.88. Thus, if one uses the GAD-7 as a screen for these multiple anxiety spectrum conditions, a midrange cutoff of 7 (sensitivity 0.80, specificity 0.76) might be more reasonable. The cutoff score to be used depends on the intended clinical use.

Funding Source

Many screening tool studies, including the PRIME-MD (and subsequently the PHQ), GAD-7, SDDS-PC, and PDI-4, were developed in conjunction with pharmaceutical companies or had researchers with pharmaceutical company relationships, raising the possibility of bias toward increased diagnoses that would lead to increased drug therapy.

Screening for One Diagnosis or More

There remains a question as to whether screening tools are more useful to assess for a single diagnosis or whether it is better to screen for multiple diagnoses at once. Even screens for one type of disorder have been found to be effective as screens for other diagnoses, thus rendering a decreased specificity. The BAI, for example, was designed to distinguish between depression and anxiety but in other studies tested positive for both; similarly, a depression scale, the Center for Epidemiologic Studies Depression Scale (CES-D), tested positive for patients with anxiety. The GAD-7 was developed to screen for generalized anxiety disorder, but it also could be used to simultaneously screen for panic disorder, social anxiety disorder, and PTSD. In addition to the effect screening multiple diagnoses might have on operating characteristics, one must also consider the time it takes to complete the screening tools, as well as the clinical capabilities to treat all the diagnoses effectively. Broader screens such as the PDI-4, a self-completed 17-item screening tool that screens for major depressive episode, generalized anxiety disorder, ADHD, and bipolar affective disorder type 1, and the My Mood Monitor checklist, which screens for depression, bipolar spectrum disorders, and anxiety disorders, may prove useful pending further validation studies.

The Elderly

Late-life depression is often underdiagnosed and subsequently undertreated. Well-studied tools in the geriatric population include the PHQ-9 and the 15-item Geriatric Depression Scale (GDS). In a study of persons 60 years or older from primary care practices, scores of greater than 5 on the GDS had a sensitivity of 0.92 and a specificity of 0.81 for major depression. In elderly primary care patients, the PHQ-9 performed comparably to the PHQ-2 and the GDS for detecting major depression. The PHQ-9 performed comparably regardless of gender or race and was somewhat better for younger elders and for those with fewer chronic illnesses. The PHQ-2 was also validated in 8000 adults older than 65 years. Overall, any of these scales perform well for screening for depression in the elderly.

Studies of anxiety screening tools tended to have younger patients, however, and there is some evidence that geriatric populations may require different screening cutoffs or instruments. In the elderly, anxiety disorders may have overlapping symptoms of chronic medical conditions, and cognitive impairment may further complicate accurate identification. In elderly patients, for example, the GAD-7 may require a lower cutoff of 5 to improve sensitivity while maintaining specificity. Alternatively, other
assessment tools such as the Geriatric Anxiety Inventory (GAI) and its 5-item short-
ened version (GAI-SF) have been validated against DSM-IV diagnostic interviews
with a sensitivity of 0.75 and specificity of 0.84 (see Table 4).58,59

**TARGETED VERSUS GENERAL SCREENING**

Should these screening tools be used for the general population, for targeted high-risk
groups, or both? Depression and anxiety disorders are of sufficient prevalence that
general primary care screening would seem reasonable, despite concerns of uncer-
tainty regarding definitive cost-benefit results. Comorbid psychiatric illness is com-
mon both concurrently as well as associated with increased likelihood of another
disorder arising later in life.60,61 Thus, there is a rationale either to screen the general
population for both depressive disorders and anxiety spectrum disorders simulta-
neously or to screen for the other if one disorder is diagnosed.

Medical conditions have been associated with both depression and anxiety. Certain
populations have demonstrated high rates of depression in association with medical
comorbidities such as cardiovascular diseases or common social risk factors such
as in veterans (see Table 3). Anxiety disorders may have higher prevalence in patients
with other medical or psychiatric diagnoses, and there is evidence of efficacy of
screening in these populations (see Table 4). For anxiety disorders especially, testing
in these populations is more problematic with regard to screening, as the somatic
symptoms may be one of the primary manifestations of an underlying anxiety or
depression spectrum disorder rather than a comorbidity—whether to call this scre-
ening may be merely semantic.

Finally, in the current Internet age, patients may complete screening tests online by
their own initiative or prompted by insurance companies or employers and bring the
results to the practitioner to review. In such settings, however, providers should clarify
whether the patient completed the tool as a screen or because of a concern regarding
symptoms.

**NOVEL AND ALTERNATIVE SCREENING MODALITIES**

In addition to traditional screening with fixed-length questionnaires given in an
office-based setting, there has been development of newer techniques (Table 5),
including computerized, adaptive testing using proprietary algorithms, screening
outside of the clinic using screening tools on the Internet, or by telephone inter-
view. Although promising, these methodologies continue to require further study
and refinement.

**RECOMMENDATIONS**

*General Recommendations*

It is essential that a screening tool be recognized as just that, a screening process for
which a positive test is not synonymous with a diagnosis but which requires additional
evaluation by a trained clinician (Box 8). Although there are no precise data to clarify
the optimal time to administer these screening tests, or in whom, preventative health
visits provide an opportune time to administer screening in the general population,
with consideration of targeted screening as other diagnoses arise. In most cases,
these tests may be self-administered. Office staff may give these screening tools to
patients with the intention to complete the test while in an appointment. Some of
the screening tools have translations in multiple languages. These tests could be
administered as part of a more comprehensive questionnaire that includes
nonpsychiatric conditions, although it is uncertain whether that strategy will affect accuracy or completion rates of the psychiatric screening tool.

**Depression Screening**

Given inconclusive data on efficacy and cost-effectiveness, official recommendations for screening for depression differ widely. Data suggest that screening is most effective in a collaborative care model (ie, integrated care with a medical doctor, case manager, and mental health specialist). Although the effectiveness of

### Box 8

**Screening for psychiatric conditions: general pearls**

1. Published properties (sensitivity, specificity) of screening tools depend on the population studied and may be different in clinical practice.

2. Positive test results are not synonymous with a diagnosis. Positive screen results must be followed by clinical assessment for a diagnosis, as only a portion of patients testing positive have a confirmed diagnosis.

3. Increasingly, these tools are able to be self-administered by patients, either in a clinical encounter or outside the office.

4. As many of these tools are widely available on the Internet, patients may find these tools themselves and bring results to the attention of their providers.

5. The optimal time for and method of screening is unknown. Preventive health visits represent an opportunity to administer screening tests.

6. Consider targeted screening depending on the condition and the populations represented in a given clinical practice.

### Table 5

**Newer methods for screening of common psychiatric conditions**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Description</th>
<th>Potential Uses</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computerized adaptive testing</td>
<td>Computer-based questions ask follow-up questions depending on the response. Duration may vary depending on assessed accuracy of diagnosis.</td>
<td>Internet-based screening in the office or at home</td>
<td>Need for computer, language capabilities, not widely available.</td>
</tr>
<tr>
<td>Telephone</td>
<td>Using same screening tools, but administered by phone</td>
<td>Outreach to patients with barriers to coming in to the office</td>
<td>Patients would still need to be seen to clarify diagnosis and to start treatment; uncertain utilization of resources.</td>
</tr>
<tr>
<td>Internet-based screening</td>
<td>Could be done by clinic, by Internet at large (eg, advocacy or nonprofit sites, but could also be from pharmaceutical companies), or by insurance companies</td>
<td>Patient completed, may be more efficient for patient to complete at home</td>
<td>Less data for real-world use and how to integrate into clinical setting.</td>
</tr>
</tbody>
</table>
screening without ancillary staff support is unclear, screening with the tools currently available requires little investment on the part of the practitioner and confers minimal immediate harm to patients themselves. Therefore, it is reasonable to screen for depression in the primary care setting, although it is best when ancillary support staff and specialist referral are available. The most readily usable screening tool is likely the PHQ-9 given its free availability, brief administration time, correlation with the DSM-IV criteria, and ability to track progress. Many current electronic systems, such as that in the Veterans Administration, use the PHQ-2 as a preliminary screen/clinical reminder, which, if positive, prompts the physician or care provider to complete a PHQ-9 or conduct a more thorough clinical interview. One may consider screening in populations such as veterans, those with postmyocardial infarction, poststroke, selected other medical conditions, and those with comorbid psychiatric illness (see Table 3). Positive diagnoses on further assessment should be assessed for suicidality, and screening for bipolar disorder should also be considered (screening for bipolar disorder specifically is not reviewed in this article) (see Box 2).

**Recommendations from major organizations**

The USPSTF recommends routine depression screening for all average-risk patients when there is sufficient staff-assisted depression care supports in place to ensure proper diagnosis, treatment, and follow-up (Table 6). In contrast, because of concerns for a high rate of false-positive diagnoses and harms of unnecessary treatment and absence of high-quality evidence for the effectiveness of screening for depression, the Canadian Task Force on Preventative Health Care (CTFPHC) revised its guidelines in 2013 and recommend against routine screening of average-risk and increased-risk individuals (although this is a weak recommendation based on very-low-quality evidence). Neither the USPSTF nor the CTFPHC make any recommendations on which screening test to use. The United Kingdom National Institute for Health and Clinical Excellence (NICE) guidelines suggest a targeted approach, screening only those individuals at risk for depression (including those with a history of depression or a chronic physical health problem with functional impairment), using the 2 PHQ-2 questions for screening.

**Anxiety Screening**

Although conclusive cost-effectiveness is lacking, given the prevalence of the disease, the available treatment, and the multitude of screening tools, it is reasonable to consider screening for anxiety disorders in the office-based setting. Owing to its ease of administration, good performance characteristics, and free distribution, the GAD-7 is a reasonable first screening tool for anxiety disorders in the primary care setting. In addition to considering use at preventive health visits, one may consider screening as an adjunctive tool in patients with depression, addiction, and unexplained somatic symptoms, with the caution that this strategy of use, while effective at identifying cases, is not well validated with respect to cost-effectiveness and patient outcomes. The ideal cutoff score for GAD-7 for generalized anxiety disorder is probably 10 or more, with a sensitivity of 0.89 and specificity of 0.82. If one’s practice has sufficient resources to treat other anxiety spectrum disorders, it may be reasonable to use a lower cutoff point of 7 to provide more sensitivity in identifying the additional conditions of panic disorder, social anxiety disorder, and PTSD (0.80 sensitivity and 0.76 specificity). If screening a general population as part of other screening questions, the GAD-2 may be more feasible to administer in the office because of its shorter length, with a follow-up GAD-7 for positive screens.
<table>
<thead>
<tr>
<th>Condition</th>
<th>Organization</th>
<th>Recommendation</th>
<th>Screening Tool</th>
<th>Strength of Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>US Preventive Services Task Force (USPSTF) (2009)</td>
<td>Routine depression screening for all average-risk patients when there is sufficient staff-assisted depression care support in place to ensure proper diagnosis, treatment, and follow-up</td>
<td>Not specified</td>
<td>Grade B: Recommended, high certainty that the net benefit is moderate or moderate certainty that the net benefit is moderate to substantial</td>
</tr>
<tr>
<td></td>
<td>Canadian Task Force on Preventative Health Care (CTFPHC) (2013)</td>
<td>Recommends against routine screening of average-risk and increased-risk individuals</td>
<td>N/A</td>
<td>Weak recommendation, very-low-quality evidence</td>
</tr>
<tr>
<td></td>
<td>United Kingdom National Institute for Health and Clinical Excellence (NICE) (2009)</td>
<td>Screen those at risk for depression (ie, history of depression, diabetes or coronary heart disease, disability, or dementia)</td>
<td>PHQ-2</td>
<td>Not specified</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td>USPSTF</td>
<td>Not addressed</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>CTFPHC</td>
<td>Not addressed</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>United Kingdom National Institute for Health and Clinical Excellence (NICE) (2011)</td>
<td>Assess in “people presenting with anxiety or significant worry,” in patients who seek care frequently who have somatic symptoms, chronic physical health problems, or “are repeatedly worrying about a wide range of different issues”</td>
<td>None specified</td>
<td>Not specified</td>
</tr>
</tbody>
</table>
Recommendations from major organizations

The NICE guidelines from the United Kingdom recommend considering the diagnosis of GAD in “people presenting with anxiety or significant worry,” and in patients who seek care frequently who have somatic symptoms, chronic physical health problems, or “are repeatedly worrying about a wide range of different issues” (see Table 6). However, these guidelines do not advocate for a particular screening method. The USPSTF, in its recommendation for screening for depression, mentions anxiety as a comorbid psychological condition that may merit increased depression screening, but there is no separate screening guideline for anxiety disorders. Guidelines will likely continue to evolve as more research into cost-effectiveness is conducted.

SUMMARY

Depression and anxiety disorders remain significant conditions in the primary care setting and in the general population. Screening tools for depression and anxiety disorders are freely available with acceptable sensitivity and specificity. Screening tools for other conditions including ADHD and PTSD also exist. Novel screening methods, including Internet-based and computerized adaptive testing, are in development and may be promising tools in the future. Despite the availability of these tools and a need to improve the mental health of patients, the utility of widespread use of screening for depression and anxiety disorders in the primary care, office-based setting is uncertain, and guidelines have reached different conclusions. The best evidence for cost-effectiveness currently is for screening of major depression as part of the collaborative care model for treatment. Targeted screening is another reasonable approach in patients with comorbid psychiatric conditions or certain medical conditions.

Despite unanswered questions, with further research, a growing literature, and increased awareness of mental health, there is every reason for optimism for the future of mental health screening.

REFERENCES


primary care: a systematic review. Washington, DC: Department of Veterans Affairs (US); 2011.


