Diagnosis

Review: questionnaires for detecting clinical depression in primary care have similar diagnostic accuracy


QUESTION: In primary care patients, what is the accuracy of questionnaires for diagnosing clinical depression?

Data sources
Studies were identified by searching Medline and a specialised registry of depression trials (1970 to July 2000).

Study selection
Published English language studies were selected if they were done in a primary care setting, administered a case-finding instrument, and obtained a criterion-based diagnosis (eg, Diagnostic and Statistical Manual of Mental Disorders, 3rd edition). The case-finding instrument had to have average literacy requirements, be scored without a calculator, have a depression specific component, and be evaluated in ≥ 1 study with ≥ 100 patients.

Data extraction
2 reviewers independently extracted data on the instrument, time frame of questions, score range, cut point, literacy level, administration time, ability of the instrument to monitor severity or response, and likelihood ratios. The quality of studies was assessed.

Main results
28 studies (25 550 screened patients) with 37 evaluations of 11 questionnaires (1 to 30 items) met the selection criteria. All questionnaires could be self administered in < 5 minutes. In 9 studies, > 50% of patients did not receive the diagnostic reference standard; 15 studies were rated as high quality. Scales for detecting major depressive disorder (MDD) did not differ (table). Heterogeneity was seen among studies assessing the following scales: Beck Depression Inventory, Center for Epidemiologic Studies Depression, Hopkins Symptom Checklist, and Zung Self-Rating Depression Scale.

Test characteristics of case-finding instruments for major depressive disorder in primary care *

<table>
<thead>
<tr>
<th>Instruments</th>
<th>Number of studies</th>
<th>Summary -LR (95% CI)</th>
<th>Summary -LR (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck Depression Inventory</td>
<td>4†</td>
<td>4.2 (1.2 to 14)</td>
<td>0.17 (0.1 to 0.3)</td>
</tr>
<tr>
<td>Center for Epidemiologic Studies Depression</td>
<td>10†</td>
<td>3.3 (2.5 to 4.4)</td>
<td>0.24 (0.2 to 0.3)</td>
</tr>
<tr>
<td>Geriatric Depression Scale</td>
<td>2</td>
<td>3.3 (2.4 to 4.7)</td>
<td>0.16 (0.1 to 0.3)</td>
</tr>
<tr>
<td>Hopkins Symptom Checklist</td>
<td>2†</td>
<td>3.2 (1.7 to 6.2)</td>
<td>0.24 (0.1 to 0.5)</td>
</tr>
<tr>
<td>Primary Care Evaluation of Mental Disorders (PRIME-MD)</td>
<td>2</td>
<td>2.7 (2.0 to 3.7)</td>
<td>0.14 (0.1 to 0.3)</td>
</tr>
<tr>
<td>PRIME-MD Patient Health Questionnaire</td>
<td>1</td>
<td>12 (8.4 to 18)</td>
<td>0.28 (0.2 to 0.5)</td>
</tr>
<tr>
<td>Symptom Driven Diagnostic System — Primary Care</td>
<td>4</td>
<td>3.5 (2.4 to 5.1)</td>
<td>0.22 (0.1 to 0.4)</td>
</tr>
<tr>
<td>Zung Self-Rating Depression Scale</td>
<td>4†</td>
<td>3.3 (1.3 to 8.1)</td>
<td>0.35 (0.2 to 0.8)</td>
</tr>
<tr>
<td>Single question</td>
<td>1</td>
<td>2.3 (1.8 to 2.9)</td>
<td>0.16 (0.0 to 0.6)</td>
</tr>
</tbody>
</table>

*Diagnostic terms defined in glossary.
†Statistically significant heterogeneity existed among studies.

Sources of funding:
San Antonio Veterans Affairs Health Services Research and Development Center of Excellence/Agency for Healthcare Quality and Research.

For correspondence:
Dr JW Williams Jr, Department of Veterans Affairs Medical Center, Durham, North Carolina, USA. jwwilliams@duke.edu

COMMENTARY

Williams et al offer a rigorous review and synthesis of the test characteristics of various depressive diagnostic survey tools derived from primary care populations. The US Preventive Services Task Force recently published a recommendation for screening patients in primary care, so this review is particularly relevant. Evidence suggests that screening for depression can improve outcomes, particularly when coupled with system changes that ensure adequate treatment and follow up.

The synthesis of literature in this study was methodologically sound. It would have been helpful to know the prevalence of MDD in each study because the test characteristics of the tools were derived in “case-finding” screening settings (ie, low prevalence settings). The derivation of test characteristics is subject to “disease spectrum bias”; thus, the test characteristics may not behave similarly in cohorts with a higher prevalence of disease (ie, people who are stressed or have multiple unexplained physical symptoms).

Health systems that plan on incorporating a systematic tool to assist in the diagnosis and monitoring of patients with MDD will benefit from the analysis by Williams et al (a useful list of web-based resources to access each tool is provided), and given that the available tools are similar in their diagnostic accuracy, the choice of which tool to use should depend on ease of use and the ability to track disease severity. It appears that the Primary Care Evaluation of Mental Disorders Patient Health Questionnaire fits these needs best. It is limited only by its inability to specifically diagnose dysthymia, a condition known to be responsive to antidepressant treatment. Busy clinicians should also note the accuracy of the Single Question instrument.

We have come a long way in developing tools that help to detect and diagnose MDD in primary care settings, but further work is needed before we can implement these tools practically and efficiently to improve depression outcomes in primary care.

Patrick G O’Malley, MD, MPH, FACP
Walter Reed Army Medical Center
Washington, DC, USA