Diagnosis

Short scales were as effective as long scales in screening for depression in older patients


QUESTION: In older patients, are short scales as effective as long scales in screening for depression?

Design
Blinded comparison of 3 Geriatric Depression Scales (GDS-4, GDS-15, and GDS-30) and the Mental Health Inventory-1 (MHI-1) scale with diagnostic criteria of research of International Classification of Diseases—10th edition (DCR-10).

Setting
A teaching hospital in the UK.

Patients
87 patients (mean age 79 y, 60% women) who were > 60 years of age and attended the day rehabilitation facility or were admitted to the medical rehabilitation wards. Patients were excluded if they had an illness, communication problems, or a score of < 6 on the 10-item Abbreviated Mental Test (AMT).

Description of tests and diagnostic standard
The GDS-30, GDS-4, 10-item AMT, and 5-item MHI were administered within 48 hours after an initial interview. Data for GDS-15 and MHI-1 were extracted from GDS-30 and MHI-5. Established cut points for diagnosis of depression were used for GDS-30, GDS-15, and GDS-4. The cut point for MHI-1 was set retrospectively. The clinical interview assessed mood and depression by using the DCR-10 (diagnostic standard).

Main outcome measures
Sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve.

Main results
17 of the 87 patients (20%) were diagnosed with depression by using the DCR-10. Sensitivity, specificity, positive and negative likelihood ratios, and ROC curve results for all tests are in the table. The 4 tests did not differ for screening of depression.

Conclusion
Short scales (Geriatric Depression Scale-4 and Mental Health Inventory-1) were comparable in sensitivity and specificity to long scales (Geriatric Depression Scale-30 and Geriatric Depression Scale-15) in screening for depression in older patients.

Operating characteristics of short and long scales to screen for depression in older patients*

<table>
<thead>
<tr>
<th>Scales</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>+LR</th>
<th>−LR</th>
<th>Area under ROC curve (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHI-1</td>
<td>88%</td>
<td>71%</td>
<td>3.08</td>
<td>0.17</td>
<td>0.88 (0.79 to 0.97)</td>
</tr>
<tr>
<td>GDS-4</td>
<td>82%</td>
<td>67%</td>
<td>2.50</td>
<td>0.26</td>
<td>0.80 (0.68 to 0.93)</td>
</tr>
<tr>
<td>GDS-15</td>
<td>82%</td>
<td>60%</td>
<td>2.06</td>
<td>0.29</td>
<td>0.82 (0.71 to 0.93)</td>
</tr>
<tr>
<td>GDS-30</td>
<td>100%</td>
<td>62%</td>
<td>2.70</td>
<td>0.00</td>
<td>0.85 (0.77 to 0.94)</td>
</tr>
</tbody>
</table>

*GDS = Geriatric Depression Scale; MHI = Mental Health Inventory; ROC = receiver operating characteristic. Other diagnostic terms defined in glossary; LR = likelihood ratio.

COMMENTARY
Depression is common, serious, and treatable, but it is under-recognised, particularly in elderly people. Societal and cultural biases often hinder the diagnosis. The study by Pomeroy et al compares 4 different screening instruments of variable length and content. The authors found that all 4 screening instruments had similar accuracy for detecting depression, and the 1-item MHI-1 had the best combination of sensitivity and specificity.

These results need further validation for 3 reasons: first, this study assessed a small inpatient sample; second, fewer than half of the patients approached were included in the study; and third, a relatively low interrater reliability (κ = 0.40) existed for MHI-1. Furthermore, the cut point for the MHI was defined retrospectively, and the item itself was not done independently and was extracted from a longer scale.

Should the clinician wait for further validation before implementing this approach to screening for depression? The answer is resounding “no.” The literature on screening for depression in general medical outpatients suggests that all of the screening instruments are relatively comparable, with sensitivity and specificity ranging from 80% to 90% similar to the findings in this study. In a typical setting, a positive test result might raise the probability of depression from 10% to 15% to 35% to 45%, and a negative test result might lower the probability of depression to < 5%. Ensuring that all patients are screened for depression regularly is more important than small changes in the precision of the screening instrument.

Scott Sherman, MD, MPH
VA Center for the Study of Healthcare Provider Behavior
Sepulveda, California, USA.