

# Diagnosis

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

Pomeroy IM, Clark CR, Philp I. *The effectiveness of very short scales for depression screening in elderly medical patients.* *Int J Geriatr Psychiatr* 2001 Mar;16:321-6.

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

Source of funding: not stated.

For correspondence: Professor I Philp, Sheffield Institute for Studies on Ageing, Northern General Hospital, Herries Road, Sheffield S5 7AU, UK.

### 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\**

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

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

### COMMENTARY

Depression is common, serious, and treatable, but it is under-recognised, particularly in elderly people. Societal and cultural biases often hinder the diagnosis.<sup>1</sup> 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 ( $\kappa = 0.40$ ) 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 a resounding “no”. The literature on screening for depression in general medical outpatients<sup>2</sup> 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.<sup>3</sup>

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

- 1 Harman JS, Reynolds CF. Removing the barriers to effective depression treatment in old age. *J Am Geriatr Soc* 2000;48:1012-3.
- 2 Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression: two questions are as good as many. *J Gen Intern Med* 1997;12:439-45.
- 3 Solberg LI, Kosen N, Oxman TE, Fischer LR, Bartels S. The need for a system in the care of depression. *J Fam Pract* 1999;48:973-9.