Which version of the geriatric depression scale is most useful in medical settings and nursing homes? Diagnostic validity meta-analysis

Mitchell AJ, Bird V, Rizzo M, Meader N

CRD summary
This review concluded that all versions of the geriatric depression scale (GDS) provided potential added value in medical settings. GDS₄/₅ was most efficient. In the absence of GDS₄/₅ data for nursing homes, GDS₁₅ may be preferred. Limitations in the methods and reporting of the review mean that these conclusions should be interpreted cautiously.

Authors’ objectives
To assess the diagnostic accuracy of different forms (long, short and ultrashort) of the Geriatric Depression Scale (GDS) in patients with and without cognitive impairment.

Searching
PubMed and EMBASE were searched from inception to October 2009. The full search strategy was reported in an appendix. Science Direct, Ingenta Select, Ovid Full text, Wiley Online Library and Web of Knowledge were searched.

Study selection
Studies that assessed the diagnostic performance of GDS long (GDS₃₀), short (GDS₁₅) and ultrashort (GDS₄/₅) for detection of depression in older people (mean age 65 years or older) with or without cognitive impairment were eligible for inclusion. Studies in medical inpatient, medical outpatient and nursing home settings were included. Included studies were required to use an adequate reference standard (psychiatric interview) to confirm diagnosis.

Where reported, mean age of study participants ranged from 66.3 to 83 years. The proportion of females ranged from zero to 84.4%. More than half of the included studies assessed major depression. The reference standard used to confirm diagnosis varied; most studies used the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders.

Two authors independently assessed studies for inclusion.

Assessment of study quality
The methodological quality of included studies was assessed using the quality assessment of diagnostic accuracy studies (QUADAS) tool.

The authors did not specify how many reviewers performed the quality assessment.

Data extraction
Sensitivity and specificity estimates for the different forms of GDS reported in each study were extracted by one reviewer.

Methods of synthesis
Pooled estimates of sensitivity and specificity, with 95% confidence intervals (CIs), were calculated for each form of the GDS. A random-effects model was used where heterogeneity was moderate to high (statistical heterogeneity test not reported). Subgroup analyses were conducted for patients with and without cognitive impairment (cognitive impairment was defined as Mini-Mental State Examination score <24) and for different study settings (in-patient, outpatient and nursing home). A Bayesian plot was constructed to show conditional post-test probabilities for all pre-test probabilities, regardless of prevalence. Area under the curve (AUC) was reported.

Results of the review
Thirty-six studies (43 data sets) were included in the review. The total number of participants was unclear. Twenty-one
studies assessed GDS\textsubscript{30}, 12 assessed GDS\textsubscript{15} and three assessed GDS\textsubscript{4/5}. The overall prevalence of late-life depression was 29.2%. Most of the included studies used an acceptable reference standard, avoided verification biases and delays between index test and reference standard and interpreted tests with access to appropriate clinical information. None of the studies reported uninterpretable results. Avoidance of incorporation bias was unclear for all studies. Blinding of test interpretation was frequently unreported. Almost half of the included studies did not recruit a representative patient spectrum.

GDS\textsubscript{30}: Overall sensitivity was 81.9% (95% CI 76.47% to 86.9%) and specificity was 77.7% (95% CI 73.0% to 82.1%). There were no significant differences in test performance between study settings or when analyses were restricted to patients with cognitive impairment or with major depression.

GDS\textsubscript{15}: Overall sensitivity was 84.3% (95% CI 79.7% to 88.4%) and specificity was 73.8% (95% CI 68.0% to 79.2%). There was no significant difference in test performance when the analysis was restricted to patients with cognitive impairment, but specificity was significantly lower in outpatients with major depression. In in-patient settings, sensitivity was 32.2% (95% CI 13.3% to 54.7%) and specificity was 69.0% (95% CI 55.4% to 81.2%). In nursing home settings, sensitivity was 86.6% (95% CI 76.1% to 94.4%) and specificity was 72.3% (95% CI 50.6% to 89.6%).

GDS\textsubscript{4/5}: Overall sensitivity was 92.5% (95% CI 85.5% to 97.4%) and specificity was 77.2% (95% CI 66.6% to 86.3%). There were insufficient data for subgroup analyses.

Bayesian analysis found no significant difference in screening or case-finding ability between GDS\textsubscript{30} and GDS\textsubscript{15}. There was a trend towards superior performance of GDS\textsubscript{15} in nursing home settings.

**Authors’ conclusions**
All versions of GDS provided potential added value in medical settings. GDS\textsubscript{4/5} was the most efficient option. In the absence of GDS\textsubscript{4/5} data for nursing home settings, GDS\textsubscript{15} may be preferred.

**CRD commentary**
The review stated a clear objective and defined appropriate inclusion criteria. Several sources were searched for relevant studies. The study selection process included measures to minimise error and/or bias; no similar measures were reported for the data extraction process. The methodological quality of included studies was assessed and reported in full. The appropriateness of meta-analyses was unclear, as the results of individual included studies were not reported and no heterogeneity assessment was reported. The evaluation of GDS\textsubscript{4/5} was based on the results of only three studies.

The authors conclusions reflect the data presented, but should be viewed cautiously given the limitations in the review methods and reporting.

**Implications of the review for practice and research**
**Practice:** The authors stated that GDS\textsubscript{15} was preferable to GDS\textsubscript{30}, for both accuracy and acceptability, in nursing home settings. They further stated that all GDS screening should be followed up with second-stage testing prior to any treatment decisions and a further assessment should be considered for patients with cognitive impairment, even in those who scored negative. A need for special care when applying self-report scales in those with substantial dementia was noted.

**Research:** The authors did not specify any recommendations for future research.

**Funding**
Not stated.

**Bibliographic details**
Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.