DIAGNOSTIC ACCURACY OF TWO BRIEF PSYCHOMETRIC MEASURES OF DEPRESSION AND ASSOCIATION WITH COGNITIVE FUNCTIONING: A STUDY INVOLVING ELDERLY MEDICAL OUTPATIENTS

Linas A. Bieliauskas, Ph.D.

Department of Veterans Affairs
Medical Center and University of Michigan,
Ann Arbor, MI 48105 USA
E-mail: linas@umich.edu

Taryn Stejskal, Ph.D.

Wellness Strategies, P.C., Indianapolis, IN USA

Brett A. Steinberg, Ph.D.

Comprehensive Neuropsychological
Services, P.C., Cheshire, CT USA

Greg J. Lamberty, Ph.D.

Department of Veterans Affairs
Medical Center, Minneapolis, MN USA

We compared the diagnostic accuracy (sensitivity and specificity) of the Geriatric Depression Scale (GDS) and the Mini-Mult Depression Scale (Mini-D) in a sample of 87 geriatric medical outpatients who were classified as depressed or non-depressed with the DSM-III-based Symptom Checklist for Major Depressive Disorders (SCMDD). In addition, we evaluated the relationship between GDS and Mini-D depression classifications and performances on three tests of the overall cognitive functioning. Although GDS and Mini-D classifications were in moderate agreement with those of the SCMDD (71%), the former measure produced more false-positive errors and the latter produced more false-negative errors. Because neither the GDS nor the Mini-D affords entirely satisfactory diagnostic accuracy, appreciation of these operating characteristics will enable practitioners to select the instrument that yields the most acceptable balance of Type I and Type II errors within their particular clinical settings. Although participants demonstrated signs of a mild cognitive compromise, no relationship was noted between depression classification and overall cognitive functioning. We believe that this finding reflects our dichotomous, rather than continuous, operationalization of "depression".

Depression is not uncommon among the elderly. According to one investigation, as many as 27% of older medical patients suffer from a depressive disorder (Reding et al., 1985). In some cases, the discovery of symptoms of depression will lead the clinician to focus chiefly upon strategies for improving the patient's mood and daily functioning. In other instances, the simultaneous discovery of signs of cognitive impairment will prompt the clinician to consider the possibility of a concurrent or underlying neurologic illness or perhaps to entertain the possibility of a mood-related decrement in cognition (so-called "pseudodementia of depression"). As a result, the patient may be referred for a more comprehensive evaluation of cognitive and emotional functioning and/or given medical treatment for the presumed neurologic disorder. In either situation, the practitioner's ability to accurately distinguish between transient or subclinical depressive symptoms and true syndromal depression will likely have important consequences for diagnosis, intervention, and prognosis.

During the past decades, the formal definition of depression has undergone few changes. According to the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, Fourth Edition, 1994), the criteria for a "major depressive episode" include a two-week history of dysphoria or anhedonia in conjunction with several additional symptoms such as diminished interest, insomnia, psychomotor retardation, or alterations in self-concept. Although many practitioners use DSM-based clinical interviews to evaluate patients for depression, such interviews are frequently time-consuming. Within the outpatient geriatric medicine clinic, brief psychometric questionnaires have become an efficient and popular alternative.

One such questionnaire, the Geriatric Depression Scale (GDS; Yesavage et al., 1983) is a thirty-item self-report measure that assesses affective symptoms (e.g., "Are you in good spirits most of the time?"), self-concept (e.g., "Do you feel pretty worthless the way you are now?"), and attitudes regarding the future (e.g., "Do you feel that your situation is hopeless?"). Examinees answer each item with "yes" or "no", and pathologic responses are summed to yield a total score which may fall within the normal (0–10), mildly depressed (11–20), or moderately to severely depressed (21–30) range. Although it has been suggested that the GDS does not assess somatic symptoms (Olin et al., 1992), this criticism seems partly unwarranted in view of the inclusion of such items as "Do you feel full of energy?" and "Do you often get restless and fidgety?" Nevertheless, the scale does not directly address disturbances of sleep, appetite, or libido. Personal experience with clinical patients suggests that the scale is susceptible to false-positive errors as a result of its emphasis on symptoms of unhappiness which may be experienced by non-clinically depressed individuals. K. B. Adams (2001) notes that a high endorsement of withdrawal and apathy items on the GDS may lead to over-identification of depression in older adults.

The Depression (D) Scale of the Minnesota Multiphasic Personality Inventory (Hathaway and McKinley, 1940) is one of the most widely used and validated measures of depression. It was based on psychiatrically diagnosed groups of patients and multiple research studies have shown its relatively linear correspondence with degrees of psychological depression (Dahlstrom et al., 1975). The MMPI D scale consists of 60 items, and a revised version of the D scale on the MMPI-2 (Butcher et al., 1989) retains 57 of these items. Lengthy test scales, however, are less desirable than shorter ones when evaluation older adults. We therefore evaluated a second psychometric screening measure of depression, the MMPI Mini-Mult Depression Scale (Mini-D; Kincannon, 1968), a twenty-item, self-report instrument that is derived from a short form of the original MMPI. Unlike the GDS, the Mini-D contains items which explicitly assess somatic symptoms (e.g., "Is your sleep fitful and disturbed?"). In addition, the scale addresses affective symptoms (e.g., "Do you wish you could be as happy as others seem to be?"), self-concept (e.g., "Are you definitely lacking in self-confidence?"), and concern regarding physical functioning (e.g., "During the last few years, have you been well most of the time?"). Owing to its empirical, as opposed to rational, development, the Mini-D featu-

res slightly less face validity than the GDS and may therefore be expected to be less vulnerable to response biases than the latter measure. Examinees answer each item with "yes" or "no", and pathologic responses are summed to yield a total score which may fall within the normal (0–6 for men, 0–9 for women) or depressed (7 or more for men, 10 or more for women) range. These cuttoff scores correspond to a T-score of 70, or two standard deviation units above the mean. We have shown that Mini-D scores of elderly patients with Parkinson's disease are closely associate with scores on the full MMPI D scale and on the Hamilton Rating Scale of Depression (Bieliauskas and Glantz, 1987), thus supporting the Mini-D as an accurate approximation of the full MMPI D scale.

Two ***goals*** guided the present research. *First*, we sought to compare the diagnostic efficacy of the GDS and the Mini-D. Given the widespread utilization of brief psychometric measures of depression and the divergent foci of these questionnaires in particular, it seemed reasonable to ask whether one instrument would be uniformly superior to the other or would be better suited to certain clinical tasks than the other. In providing such information, we hope to aid practitioners in making knowledgeable choices regarding their measurement tools. *Second*, we sought to add new data to the controversial issue of the association between depression and cognition. Given the diagnostic and therapeutic complexities that characterize cases involving both psychiatric and neurologic disorders, there appears to be an urgent need to clarify the relative contributions of, and interactions between, these factors. Although a review of the evidence that bears on this question is beyond the scope of this report, we note that some investigators have observed an inverse relationship between cognitive functioning and depression (Lichtenberg et al., 1995; Norris et al., 1995) while other researchers, including our group, have failed to find such a relationship (Bieliauskas, 1993; Bieliauskas et al., 1991; Lamberty and Bieliauskas, 1993; Nusbaum et al., 1995). It is quite possible that these divergent results reflect different operationalizations of "depression". In some cases, depression, as measured by psychometric screening instruments, has been treated as a continuous variable and correlated with scores on measures of neurocognitive abilities (Lichtenberg et al., 1995; Nusbaum et al., 1995). We believe that this approach is unsatisfactory because it obscures the meaningful distinction between depression as a formal clinical syndrome (as defined in the DSMs) and subclinical depressive symptoms which represent normal variability in affective experience within the general population. It must be remembered that in the case of all scales measuring depression, there is a "normal" range, i.e. it is quite expectable that most individuals completing any scale will positively endorse some items, and it is only when they endorsed a significant number of items that they would be considered to score in the "depressed" range. In those cases in which quantitative differences in the severity of depression have been specified only after examinees have crossed the threshold for syndromal depression, the relationship between mood disorder and cognitive disorder has essentially vanished relationship (Bieliauskas, 1993; Bieliauskas et al., 1991; Lamberty and Bieliauskas, 1993; Nusbaum et al., 1995).

As is often the case in geriatric clinical practice, participants in the present study came to the attention of their physicians as a result of subjective and/or objective changes in cognition. As part of their diagnostic workups, patients were referred for a comprehensive neuropsychologic examination in order to clarify the nature and extent of such changes.

Method

Participants were 87 geriatric medical outpatients (23 male, 64 female; mean age = 77.4 years) who were evaluated through the Neuropsychology Division at the University of Michigan Medical Center. Their evaluations involved administration of standardized tests of mental status, general intellectual functioning, specific neurocognitive abilities (attention, language, visuospatial functioning, memory, and executive functioning), and emotional functioning. At the time of their examinations, no participants carried a diagnosis of depression.

Each patient completed an orally-administered form of the GDS and the Mini-D. Because patients' diagnostic workups did not include a psychiatric evaluation, they also completed an oral form of the DSM-III-based Symptom Checklist for Major Depressive Disorders (SCMDD; Kashani et al., 1985). The latter measure is a twenty-three-item, self-report instrument that assesses nine domains of depressive symptoms (dysphoria, loss of interest in daily activities, and irritability; disturbance of appetite; disturbance of sleep; fatigability; psychomotor agitation and retardation; anhedonia; feelings of guilt and worthlessness; impaired concentration and decision-making ability; and suicidal ideation). As is the case for the GDS and the Mini-D, examinees answer each item with "yes" or "no", and a diagnosis of depression is considered appropriate when patients endorse at least one symptom from the first domain and at least one symptom from four of the remaining eight domains. Because this measure essentially reproduces the official DSM-III diagnostic criteria for major depressive episode within a brief question-and-answer format, we believe that it represents a suitable standard of comparison for evaluating the efficacy of the GDS and Mini-D (Lamberty et al., 1994).

The initial set of data analyses involved comparison of the demographic characteristics of groups of participants who were classified as depressed and non-depressed by each of the three depression measures. Using the groups that resulted from SCMDD depression classifications, the second set of analyses entailed calculation of the sensitivity (the likelihood of diagnosing a disorder when it exists) and the specificity (the likelihood of not diagnosing a disorder when it does not exist) of the GDS and the Mini-D and comparison of the classification rates of the latter measures (Mausner and Kramer, 1985). The final set of data analyses involved examination of the relationship between depression classifications and overall neuropsychologic functioning. For the purpose of these latter analyses, "overall neuropsychologic functioning" was indexed by total scores on the Mini-Mental State Examination (MMSE), memory quotient (MQ) scores on the Wechsler Memory Scale (WMS; Wechsler, 1945) and education-corrected total scores on the Controlled Oral Word Association Test (COWA; Benton and Hamsher, 1989).

Results

Inspection of the data in Table 1 indicates that the published cutoff scores for the GDS, the Mini-D, and the SCMDD produced grossly comparable groups of depressed and non-depressed participants. No significant differences were noted for mean age (middle to late seventies) or for mean level of education (approximately twelve years). In addition, the proportions of men and women in the depressed and non-depressed groups were similar for each measure (17–23% male for the depressed groups and 28–31% male for the non-depressed groups), and these proportions were generally consistent with the overall male-to-female ratio in the sample (26% male).

Although the GDS classified 34% of the sample as depressed, the Mini-D classified only 14% of the sample as depressed. Because the SCMDD resulted in classification of 33% of the sample as depressed, it may be tempting to conclude that the GDS afforded greater diagnostic accuracy than the Mini-D. Examination of Table 2 demonstrates, however, that these measures possess different classificatory strengths and weaknesses. On the one hand, the GDS offered modest sensitivity (59%) and moderate specificity (78%). On the other hand, the Mini-D offered limited sensitivity (28%) but considerable specificity (93%). The overall agreement between each of these instruments and the criterion SCMDD was moderate (71%), and chi-square tests revealed significant associations between the GDS and the SCMDD (χ^2 [1, N = 87] = 11.22, $p < .001$) and between the Mini-D

Table 1. ***Demographic characteristics of depressed and non-depressed participants as classified by the GDS, Mini-D, and SCMDD***

Measure of depression	Classification*	
	Depressed	Non-depressed
GDS		
N	30	57
Male-to-female ratio	7:23	16:41
Mean age	77.0 (6.6)	77.5 (6.7)
Mean years of education	12.1 (3.5)	13.3 (3.7)
Mini-D		
N	12	75
Male-to-female ratio	2:10	21:54
Mean age	77.8 (6.5)	77.3 (6.7)
Mean years of education	11.5 (3.0)	13.1 (3.7)
SCMDD		
N	29	58
Male-to-female ratio	5:24	18:40
Mean age	75.8 (6.1)	78.1 (6.8)
Mean years of education	12.8 (3.8)	13.0 (3.6)

Note: GDS = Geriatric Depression Scale; Mini-D = Mini-Mult Depression Scale; SCMDD = Symptom Checklist for Major Depressive Disorders (DSM-III). Standard deviations appear in parentheses.

* As determined by published cutoff scores (see text for details).

Table 2. ***Comparison of GDS and Mini-D depression classifications using SCMDD depression classification as the criterion***

Measure of depression	SCMDD classification*	
	Depressed	Non-depressed
GDS*		
Depressed	17	13
Non-depressed	12	45
Mini-D*		
Depressed	8	4
Non-depressed	21	54

Note: GDS = Geriatric Depression Scale; SCMDD = Symptom Checklist for Major Depressive Disorders (DSM-III). Standard deviations appear in parentheses.

* As determined by published cutoff scores (see text for details).

Table 3. ***Relationship between depression classifications and indices of overall neuropsychologic functioning for the GDS, Mini-D, and SCMDD***

Measure of depression	Classification*	
	Depressed	Non-depressed
GDS		
MMSE total score	20.7 (4.9)	21.3 (4.9)
WMS MQ	80.8 (15.0)	83.7 (13.3)
COWA education-corrected score	21.4 (13.8)	24.0 (13.2)
Mini-D		
MMSE total score	21.3 (5.3)	21.1 (4.9)
WMS MQ	84.0 (15.3)	82.4 (13.8)
COWA education-corrected score	23.2 (15.9)	24.0 (13.2)
SCMDD		
MMSE total score	20.9 (5.1)	21.2 (4.8)
WMS MQ	83.9 (14.8)	82.1 (13.6)
COWA education-corrected score	21.6 (13.1)	23.8 (13.6)

Note: GDS = Geriatric Depression Scale; Mini-D = Mini-Mult Depression Scale; SCMDD = Symptom Checklist for Major Depressive Disorders (DSM-III). Standard deviations appear in parentheses.

* As determined by published cutoff scores (see text for details).

and the SCMDD (χ^2 [1, N = 87] = 6.96, p = .008).

Table 3 presents data concerning the relationship between depression classifications and overall cognitive functioning. These data indicate that depressed and non-depressed participants exhibited mildly compromised mental status (using the traditional MMSE cutoff score of 24; Folstein et al., 1975), low average overall memory functioning, and borderline to mildly defective phonemically-based word list generation. No significant differences were found between the depressed and non-depressed groups on these three neuropsychologic tests.

Discussion

Fully one-third of our sample of elderly medical outpatients evinced a clinically significant mood disorder as measured by the DSM-III-based SCMDD. This prevalence rate is generally consistent with one previously published rate of 27% (Reding et al., 1985), and it provides additional support for the proposal that depression is an important problem among older individuals. Although the GDS and the Mini-D offered a moderate degree of overall classificatory agreement with the SCMDD (71%), the GDS was associated with a three-fold increase in rate of false-positives relative to the Mini-D. This finding is consistent with the aforementioned observation that the GDS tends to overdiagnose depression in elderly respondents. Although the Mini-D afforded more specificity than the GDS, this benefit was associated with a sizable decrease in sensitivity. Because neither instrument provides entirely satisfactory diagnostic accuracy, appreciation of these operating characteristics will enable geriatric medical practitioners to select the measure that yields the most acceptable balance of Type I and Type II errors within their particular clinical setting. For example, if the goal of assessment is to identify patients who will benefit from more extensive psychiatric or psychologic evaluation, then the GDS would be a reasonable choice. Although a number of non-depressed patients will undergo such evaluation, this is a rather small and benign price to pay for identifying additional patients who are depressed and in need of pharmacologic and/or psychotherapeutic management.

Nevertheless, it has been pointed out that overdiagnosis of depression may result in unnecessary adverse effects of psychotropic medications (Thorpe, 2009). Therefore, if the goal of assessment is to decide whether or not to initiate antidepressant therapy within the primary care setting, then the Mini-D may be the more prudent choice. To be sure, contemporary agents (including selective serotonin reuptake inhibitors) offer more favorable side-effect profiles than do older medications such as monoamine oxidase inhibitors. Nevertheless, the former may produce memory impairment and other anticholinergic effects in drug-sensitive geriatric patients. Because individuals who are classified as depressed by the Mini-D are indeed likely to be suffering from a major depressive episode, clinicians may be more confident when using this instrument that they are prescribing agents for bona fide mood disorders. It should be noted, however, that psychometric test data – including those from the SCMDD – are intended to be buttressed by corroborating evidence when deciding whether to assign a psychiatric

diagnosis. The absence of such evidence in the present study should be taken into consideration as the reader interprets the foregoing findings regarding the accuracy of the GDS and the Mini-D.

Consistent with several previous studies (Bieliauskas, 1993; Bieliauskas et al., 1991; Lamberty and Bieliauskas, 1993; Nusbaum et al., 1995), the present investigation failed to support the hypothesis that depression is associated with cognitive impairment in the elderly *when depression is operationalized as a dichotomous (rather than continuous) variable*. Although the SCMDD data suggested that only one-third of the sample was depressed, group means for the three measures of overall cognitive functioning of our sample were mildly compromised. It is theoretically possible that the cognitive dysfunction of the depressed patients resulted from a mood disorder and that the cognitive dysfunction of the non-depressed patients resulted from other factors, but this seems quite improbable (especially in light of the fact that depressed and non-depressed participants' MMSE and COWA performances were comparable). More reasonable, we believe, is the proposal that concurrent depression and overall cognitive impairment, as defined herein, represented either coincident findings or alternate expressions of a single pathologic process.

Note: The GDS, Mini-D, and SCMDD are readily accessible from the original references and can be translated into different languages for country-specific use.

REFERENCES

Adams K. B. Depressive symptoms, depletion, or developmental change: Withdrawal, apathy, and lack of vigor in the Geriatric Depression Scale // The Gerontologist. 2001, vol. 41, p. 768–777.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author, 1994.

Benton A. L., Hamsher K. de S. Multilingual aphasia examination. Iowa City, IA: AJA Associates, 1989.

Bieliauskas L. A. Depressed or not depressed? That is the question // Journal of Clinical and Experimental Neuropsychology. 1993, vol. 15, p. 119–134.

Bieliauskas L. A., Glantz R H. Use of the Mini-Mult D scale in patients with Parkinson's disease // Journal of Consulting and Clinical Psychology. 1987, vol. 55, p. 437–438.

Bieliauskas L. A., Lamberty G., Boczar J. Lack of depression effects on cognitive functions in the elderly (abstract) // Journal of Clinical and Experimental Neuropsychology. 1991, vol. 13, p. 433.

Butcher J. N., Dahlstrom W. G., Graham J. R., Tellegen A., Kaemmer B. Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring. Minneapolis: University of Minnesota Press, 1989.

Folstein M. F., Folstein S. E., McHugh P. R. Mini-mental state // Journal of Psychiatric Research. 1975, vol. 12, p. 189–198.

Dahlstrom W. G., Welsh G. S., Dahlstrom L. An MMPI handbook (Vol I.). Minneapolis: University of Minnesota Press, 1975.

Hathaway S. R., McKinley J. C. A multiphasic personality schedule (Minnesota): I, Construction of the schedule // Journal of Psychology. 1940, vol. 10, p. 249–254.

Kashani J., McKnew D., Cytryn L. Symptom checklist for major depressive disorders // Psychopharmacology Bulletin. 1985, vol. 21, p. 957–958.

Kincannon J. C. Prediction of the standard MMPI scale scores from 71 items: The mini-mult. // Journal of Consulting and Clinical Psychology. 1968, vol. 32, p. 319–325.

Lamberty G. J., Bieliauskas L. A. Distinguishing between depression and dementia in the elderly: A review of neuropsychological findings // Archives of Clinical Neuropsychology. 1993, vol. 8, p. 149–170.

Lamberty G. J., Bieliauskas L. A., Holt C. S. Depressive symptom covariation in geriatric clinic patients // The Clinical Gerontologist. 1994, vol. 15, p. 15–27.

Lichtenberg G. J., Ross T., Millis S. R., Manning C. A. The relationship between depression and cognition in older adults: A cross-validation study // Journal of Gerontology: Psychological Sciences. 1995, vol. 50, p. 25–32.

Mausner J. S., Kramer S. Mausner and Bahn epidemiology: An introductory text. Philadelphia, PA: WB Saunders Company, 1985.

Norris M. P., Blankenship-Reuter L., Snow-Turek A. L., Finch J. Influence of depression on verbal fluency performance // Aging and Cognition. 1995, vol. 2, p. 206–215.

Nusbaum P. D., Kaszniak A. W., Allender J., Rapcsak S. Depression and cognitive decline in the elderly: A follow-up study // The Clinical Neuropsychologist. 1995, vol. 9, p. 101–111.

Olin J. T., Schneider L. S., Eaton E. M. et al. The Geriatric Depression Scale and the Beck Depression Inventory as screening instruments in an older adult outpatient population // Psychological Assessment. 1992, vol. 4, p. 190–192.

Reding M., Haycox J., Blass, J. Depression in patients referred to a dementia clinic: A three-year prospective study // Archives of Neurology. 1985, vol. 42, p. 894–896.

Thorpe L. Depression vs. dementia: How do we assess? // The Canadian Review of Alzhiemer's Disease and other dementias. 2009, vol. 12, p. 17–21.

Wechsler D. A standardized memory scale for clinical use // Journal of Psychology. 1945, vol. 19, p. 87–95.

Yesavage. J. A., Brink T. L., Rose T. L. et al. Development and validation of a geriatric depression screening scale: A preliminary report // Journal of Psychiatric Research. 1983, vol. 17, p. 37–49.

DVIEJŲ TRUMPŲ PSICHOMETRINIŲ DEPRESIJOS PAŽINIMO FUNKCIJOS MASTŲ DIAGNOSTINIS TIKSLUMAS: VYRESNIO AMŽIAUS AMBULATORINIŲ PACIENTŲ STUDIJA

Linas A. Bieliauskas, Taryn Stejskal, Brett A. Steinberg, Greg J. Lamberty

S a n t r a u k a

Lyginome Senyvųjų depresijos skalės (*Geriatric Depression Scale*) ir Mažosios daugkartinės depresijos skalės (*Mini-Mult Depression Scale*) diagnostinį tikslumą (jautrumą ir tikslų apibrėžimą). Tyrime dalyvavo 87 senyvi gydomi pacientai, kai kurie jų buvo depresyvūs, o kai kurie nedepresyvūs (buvo naudojamasi DSM-III simptomų sąrašu, nurodančiu didžiosios depresijos sutrikimą (SCMDD)). Be to, įvertinome Senyvųjų depresijos skalės ir Mažosios daugkartinės depresijos skalės klasifikacijų santykį naudodami tris bendro pažinimo funkcijos testus. Nors Senyvųjų depresijos skalė ir Mažoji daugkartinės depresijos skalė pagal DSM-III sąrašą rodė vidutinį lygumą (71 %), pirmoji grupė darė daugiau pozityvių klaidų, o antroji – daugiau negatyvių klaidų. Kadangi nei Senyvųjų depresijos skalės, nei Mažosios daugkartinės depresijos skalbės diagnostinis tikslumas nėra pakankamas, dėmesio atkreipimas į šias charakteristikas turėtų paskatinti naudojančiuosius jas savo klinikiniame darbe pasirinkti priemonę, kurios 1-o ir 2-o tipo klaidų balansas geriausias. Nors pacientų pažinimo sutrikimo požymiai buvo neryškūs, jokio santykio tarp depresijos klasifikacijos ir bendro pažinimo funkcionavimo neaptikome. Manome, kad šis rezultatas rodo, jog, vartojant „depresijos" sąvoką, yra dichotomija, o ne vientisumas.

Įteikta 2010-10-16
Pataisytas straipsnis įteiktas 2011-03-03