J Psychopathol Behav Assess (2009) 31:412–416
DOI 10.1007/s10862-008-9123-z

Internal Consistency and Test-Retest Stability of the Geriatric Depression Scale-Short Form in African American Older Adults

Otto Pedraza · Vonetta M. Dotson · Floyd B. Willis · Neill R. Graff-Radford · John A. Lucas

Published online: 8 January 2009

Abstract The Geriatric Depression Scale (GDS) is one of the most widely used self-rated mood questionnaires for older adults. It is highly correlated with clinical diagnoses of depression and has demonstrated validity across different patient populations. However, the reliability of the GDS among African American older adults remains to be firmly established. In a baseline sample of 401 African American adults age 51 and over, the GDS-15 item short form demonstrates good internal consistency (KR20=.71). Stability over a 15-month interval in a retest sample of 51 adults is deemed adequate (r=.68). These findings support the use of the GDS-15 item short form as a reliable mood questionnaire among African American older adults.

Keywords Geriatric depression scale · GDS · Depression · African American · Reliability

O. Pedraza (✉) · J. A. Lucas
Department of Psychiatry and Psychology, Mayo Clinic,
4500 San Pablo Road,
Jacksonville, FL, USA 32224
e-mail: otto.pedraza@mayo.edu

V. M. Dotson
Laboratory of Personality and Cognition,
Intramural Research Program, National Institute on Aging,
National Institutes of Health,
Baltimore, MD, USA

F. B. Willis
Department of Family Medicine, Mayo Clinic,
Jacksonville, FL, USA

N. R. Graff-Radford
Department of Neurology, Mayo Clinic,
Jacksonville, FL, USA

As the population of the United States ages and becomes more racially heterogeneous (He et al. 2005), there is a vital need for reliable tools to assess psychological symptoms in ethnically diverse elders. Precise measurement of depressive symptoms is particularly important given the significant morbidity and mortality associated with clinical depression and subthreshold depressive symptoms among older adults (Chapman and Perry 2008). Moreover, the incidence of clinically significant depressive symptoms increases steadily with advancing age (Blazer 2003; Lavretsky and Kumar 2003) and prevalence rates ranging from 4% to 36% in the community (Beck and Koenig 1996; Blazer 1994; Callahan et al. 1994; Judd et al. 1994; Lyness et al. 1999) and 25% to 64% in medical settings (Koenig et al. 1997; Rapp et al. 1991; Rosen et al. 1997) have been reported.

Among African American adults, it remains unclear whether the incidence and prevalence rates of depression are comparable to those of White adults. A few studies report similar (Blazer et al. 1998; Mills et al. 2004) or lower (Gallo et al. 1998) symptoms of depression, but other reports suggest a higher rate of depressive symptomatology in older African Americans (Boutin-Foster 2008; Skarupski et al. 2005). For instance, in a population-based study of 7,690 older adults between the ages of 54 and 65, Dunlop et al. (2003) found that the unadjusted prevalence of depression was greater in African Americans than Whites, but that the risk ratio was strongly mediated by sociodemographic, health, and economic factors.

Furthermore, numerous studies have documented variations in the clinical presentation of late-life depression in African Americans and other ethnic minority groups (Baker et al. 1996; Blazer et al. 1998; Gallo et al. 1998; Kirmayer 2001). For example, African Americans tend to report less dysphoria (Gallo et al. 1998), but more hopelessness and somatic symptoms than Whites (Baker et al., 1996; Blazer

et al. 1998). African Americans may conceptualize depression in a different manner, have different health beliefs about depression, or attach greater stigma to mental illness, all of which affect responses on self-rated depression scales (Mills et al. 2004). These variations in the clinical presentation of depression underscore the need for precise measurement instruments within this demographic group.

The Geriatric Depression Scale (GDS; Brink et al. 1982; Yesavage et al. 1983) is one of the most widely used self-rated depression questionnaires for older adults. The GDS was specifically developed to screen for geriatric depression, and as such, it was designed to avoid somatic symptoms that complicate the differential diagnosis when comorbid medical conditions are present (Blazer 2002; Yesavage et al. 1983). It is highly correlated with clinical diagnoses of depression (Wall et al. 1999) and has demonstrated validity across different patient populations, including hospital inpatients, outpatients, nursing home residents, and psychiatric patients (Ertan et al. 2005; Friedman et al. 2005; McGivney et al. 1994; Stiles and McGarrahan 1998; Wancata et al. 2006). Shorter forms of the GDS have been developed (Sheikh and Yesavage 1986; Yesavage 1988), and the GDS has been translated into many languages and validated in a variety of different countries.

In a review of 338 studies, Kieffer and Reese (2002) found 98 peer-reviewed publications that reported reliability estimates for the GDS. Thirty-five of these studies reported more than one type of reliability coefficient (e.g., internal consistency, test-retest) and 40 studies specifically examined the 15-item version of the GDS. Across all studies, the mean internal consistency coefficient was .85 and the test-retest coefficient was .83. Study sample size, standard deviations, and number of items on the instrument (i.e., 30, 15, 10, or 4) greatly influenced the obtained reliability coefficients.

However, the reliability of the GDS remains to be firmly established in African American adults. This void is particularly concerning given the aforementioned differences in prevalence and clinical presentation among older African Americans compared to Caucasian adults. Kurlowicz et al. (2005) examined the GDS-30 item version among 150 African American adults evaluated at an academic outpatient rehabilitation program and found that 30% of the sample was experiencing depression when using the recommended cutoff score of 11 points or higher. Unfortunately, the psychometric properties of the instrument in this demographic group were not analyzed. Consequently, the goal of the present study is to investigate the reliability of the GDS in a community sample of African American older adults in order to enhance the instrument's known psychometric properties.

Method

Participants

Participants included 401 self-identified African American adults who completed the Geriatric Depression Scale-15 item short form at baseline as part of the Mayo Older African Americans Normative Studies (MOAANS) protocol. Study criteria and recruitment protocol for the MOAANS protocol have been described previously (Lucas et al. 2005). Participants consisted of cognitively normal, community dwelling, independently functioning individuals examined by their primary physician within 1 year of study entry who met the following selection criteria: (1) Normal cognitive functioning based on self-report, informant report, and physician report; (2) normal cognitive capacity to perform independent activities of daily living based on informant report; (3) no active or uncontrolled central nervous system, systemic, or psychiatric condition that would adversely affect cognition, based on physician report; (4) no use of psychoactive medications in amounts that would be expected to compromise cognition or for reasons indicating a primary neurologic disease or psychiatric illness; and (5) no prior history of dementia, stroke, movement disorder, multiple sclerosis, brain tumor, seizures, severe head trauma, schizophrenia, bipolar mood disorder, or major depression. Patients with histories of other conditions having the potential to affect cognition (e.g., heart attack, mild head injury) were allowed to participate if no evidence of residual cognitive dysfunction was evident on general medical evaluation and their physician judged the condition to be no longer active. Patients with active, chronic medical disease (e.g., diabetes, hypertension, thyroid dysfunction) were included if their physician judged that the condition was adequately controlled and not causing cognitive compromise. It should be noted that the above criteria did not preclude enrollment of individuals experiencing symptoms of depression, but only those with a prior history of clinically diagnosed major depression or whose depression at the time of assessment was judged to be sufficiently severe that it interfered with their cognition. A convenience sample of 51 participants was available and retested within a 15-month period and provide the GDS scores for stability calculations.

All data were obtained in full compliance with study protocols approved by the Mayo Clinic Institutional Review Board.

Measure

The Geriatric Depression Scale-short form comprises 15 items presented in a forced-choice yes/no response format. The GDS was administered as part of a lengthier

neuropsychological evaluation within the MOAANS protocol. A score of 6/15 or higher prompted a message from the research coordinator to one of three investigators (J.A.L, N. R.G., or F.B.W.) for further assessment of the participant's mood. If deemed clinically necessary, the patient was then offered consultation with a mental health professional for clinical evaluation and consideration of treatment options. Thirty-three out of the 401 baseline participants obtained a GDS score of six or higher and were contacted for subsequent assessment.

Statistical Analyses

Internal consistency was estimated with the Kuder-Richardson 20 (KR20), a case of Cronbach's alpha coefficient for dichotomous items. Test-retest stability was estimated as the bivariate correlation between the baseline and follow-up GDS scores. The relationship between demographic variables and GDS scores was analyzed using t-tests or bivariate correlations. All statistical analyses were performed using SPSS, version 14.0.

Results

Sample characteristics are presented in Table 1 and suggest broad representation of African American older adults in age (range 51–94) and education (range 0–20). The GDS score at baseline was inversely correlated with age (r=−.17, p<.001) and education (r=−.14, p<.01), and these associations remained significant in partial correlation analyses. Mean GDS score in men (M=2.4, S.D.=2.4) and women (M=2.1, S.D.=2.2) were not significantly different (t(399)=1.1, p=.26). The GDS showed good internal consistency (KR20=.71) in this sample.

The GDS was readministered to 51 participants within 15 months after their baseline session (Table 1). The range of GDS scores was comparable between the two assessment periods and there was no significant interval difference in mean GDS scores (t(50)=−.92, p=.36). Test-retest stability was r=.68.

Table 1 Sample characteristics and Geriatric Depression Scale (GDS) reliability coefficients

	Mean	S.D.	Range
Baseline sample (n=401)			
Age	70.4	7.8	51–94
Education (years)	12.2	3.4	0–20
GDS total score	2.2	2.2	0–11
Internal consistency r=.71			
Follow-up sample (n=51)			
Retest interval (days)	393	35.7	337–455
GDS total score, time 1	1.2	1.9	0–10
GDS total score, time 2	1.4	2.2	0–11
Test-retest stability r=.68			

Discussion

Despite numerous clinical and cross-cultural investigations on the psychometric properties of the GDS, its reliability among African American older adults remains to be firmly established. The current brief report aims to address this gap by investigating the internal consistency and test-retest stability of the GDS-15 short form in cognitively-normal, community-dwelling adults. The results show that the GDS has good internal consistency (r=.71) and demonstrates adequate test-retest stability (r=.68) within a 15-month interval period. These coefficients are lower than the mean reliability estimate of 0.82 reported by Kieffer and Reese (2002) for the 15 item version of the GDS. That review, however, found a considerable range of reliability coefficients. Across 117 studies in which reliability coefficients were not reported but could be estimated, the authors found a mean reliability estimate of 0.76 and a range from 0.11 to 0.99.

An important consideration is that calculation of test-restest stability may not be theoretically justified in questionnaires designed to assess reasonably unstable constructs. Indeed, there is no compelling argument for the use of stability coefficients to establish reliability in a mood questionnaire because the result may confound stability of the construct with the instrument's insensitivity to detect changes in the construct. Nonetheless, the data summarized by Kieffer and Reese (2002) show that approximately one-fourth of all studies exploring the reliability of the GDS have used a test-retest coefficient to gauge reliability. Stability results are presented in that context, but taking into consideration the qualifications noted earlier and underscoring that the stability coefficient obtained in the current study should be interpreted with caution.

There are several strengths to the current investigation. First, reliability coefficients were calculated using a relatively large sample of 401 community-dwelling adults with broad representation of age and education. Second, participants with medical conditions were not excluded unless those conditions contributed to cognitive impairment. This is important given the known relationship between health status and mood functioning. Third, although patients with a clinical diagnosis of major depression were excluded from the parent normative study, GDS scores in this sample ranged from 0 to 11, suggesting that there was minimal restriction of the possible range of scores. Indeed, 55 participants endorsed a minimum of five items, consistent with the cutoff criteria reflecting mildly elevated symptoms of depression (Brink et al. 1982; Sheikh

and Yesavage 1986). It could be reasonably argued, however, that the absence of a clinically diagnosed group of patients with major depression or other mood disorders represents a limiting factor insofar as the low rate of item endorsement on the GDS may temper internal consistency. Overall, these findings support the GDS-15 item version as a reliable instrument for the assessment of depressive symptoms in African American older adults.

Acknowledgments We want to extend our gratitude to Francine C. Parfitt, Tracy L. Kendall, Sylvia B. Stewart, Kelly R. Smith at the Mayo Memory Disorders Clinic for assistance with aspects of data collection and preparation. This study was supported in part by a grant from the Robert and Clarice Smith Fellowship Program, a Supplement grant from the National Institute on Aging (AG016574-07S1), and the Intramural Research Program of the NIH, National Institute on Aging. The study authors do not have any sources, financial or otherwise, that could result in a conflict of interest pertaining to this manuscript.

References

Baker, F. M., Okwumabua, J., Philipose, V., & Wong, S. (1996). Screening African-American elderly for the presence of depressive symptoms: a preliminary investigation. *Journal of Geriatric Psychiatry and Neurology*, *9*(3), 127–132.

Beck, D. A., & Koenig, H. G. (1996). Minor depression: a review of the literature. *International Journal of Psychiatry in Medicine*, *26* (2), 177–209.

Blazer, D. G. (1994). Epidemiology of late-life depression. In L. S. Schneider, C. F. Reynolds, B. D. Lebowitz, & A. Friedhoff (Eds.), *Diagnosis and treatment of late life depression: Results of the NIH consensus development conference* (pp. 9–21). Washington, DC: American Psychiatric.

Blazer, D. G. (2002). Symptoms and signs. In *Depression in late life* (pp. 35–57). New York: Springer.

Blazer, D. G. (2003). Depression in late life: review and commentary. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *58*(3), 249–265.

Blazer, D. G., Landerman, L. R., Hays, J. C., Simonsick, E. M., & Saunders, W. B. (1998). Symptoms of depression among community-dwelling elderly African-American and white older adults. *Psychological Medicine*, *28*(6), 1311–1320.

Boutin-Foster, C. (2008). An item-level analysis of the Center for Epidemiologic Studies Depression Scale (CES-D) by race and ethnicity in patients with coronary artery disease. *International Journal of Geriatric Psychiatry*, *23*(10), 1034–1039.

Brink, T. L., Yesavage, J. A., Lum, O., Heersema, P. H., Adey, M., & Rose, T. L. (1982). Screening tests for geriatric depression. *Clinical Gerontologist*, *1*, 37–43.

Callahan, C. M., Hui, S. L., Nienaber, N. A., Musick, B. S., & Tierney, W. M. (1994). Longitudinal study of depression and health services use among elderly primary care patients. *Journal of the American Geriatric Society*, *42*(8), 833–838.

Chapman, D. P., & Perry, G. S. (2008). Depression as a major component of public health for older adults. *Preventing Chronic Disease*, *5*(1), A22.

Dunlop, D. D., Song, J., Lyons, J. S., Manheim, L. M., & Chang, R. W. (2003). Racial/ethnic differences in rates of depression among preretirement adults. *American Journal of Public Health*, *93*(11), 1945–1952.

Ertan, F. S., Ertan, T., Kiziltan, G., & Uygucgil, H. (2005). Reliability and validity of the Geriatric Depression Scale in depression in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(10), 1445–1447.

Friedman, B., Heisel, M. J., & Delavan, R. L. (2005). Psychometric properties of the 15-item geriatric depression scale in functionally impaired, cognitively intact, community-dwelling elderly primary care patients. *Journal of the American Geriatric Society*, *53*(9), 1570–1576.

Gallo, J. J., Cooper-Patrick, L., & Lesikar, S. (1998). Depressive symptoms of whites and African Americans aged 60 years and older. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *53*(5), P277–286.

He, W., Sengupta, M., Velkoff, V. A., & DeBarros, K. A. (2005). *65+ in the United States: 2005 (U.S. Bureau of Census, Brief P23–209)*. Washington, DC: US Government Printing Office.

Judd, L. L., Rapaport, M. H., Paulus, M. P., & Brown, J. L. (1994). Subsyndromal symptomatic depression: a new mood disorder? *Journal of Clinical Psychiatry*, *55*(Suppl), 18–28.

Kieffer, K. M., & Reese, R. J. (2002). A reliability generalization study of the Geriatric Depression Scale. *Educational and Psychological Measurement*, *62*(6), 969–994.

Kirmayer, L. J. (2001). Cultural variations in the clinical presentation of depression and anxiety: implications for diagnosis and treatment. *Journal of Clinical Psychiatry*, *62*(Suppl 13), 22–28 discussion 29–30.

Koenig, H. G., George, L. K., Peterson, B. L., & Pieper, C. F. (1997). Depression in medically ill hospitalized older adults: prevalence, characteristics, and course of symptoms according to six diagnostic schemes. *American Journal of Psychiatry*, *154*(10), 1376–1383.

Kurlowicz, L. H., Outlaw, F. H., Ratcliffe, S. J., & Evans, L. K. (2005). An exploratory study of depression among older African American users of an academic outpatient rehabilitation program. *Archives of Psychiatric Nursing*, *19*(1), 3–9.

Lavretsky, H., & Kumar, A. (2003). Practical geriatrics: clinically significant nonmajor geriatric depression. *Psychiatr Serv*, *54*(3), 297–299.

Lucas, J. A., Ivnik, R. J., Smith, G. E., Ferman, T. J., Willis, F. B., Petersen, R. C., et al. (2005). Mayo's older African American normative studies: Norms for Boston naming test, controlled oral word association, category fluency, animal naming, token test, WRAT-3 reading, trail making test, stroop test, and judgment of line orientation. *The Clinical Neuropsychologist*, *19*, 243–269.

Lyness, J. M., King, D. A., Cox, C., Yoediono, Z., & Caine, E. D. (1999). The importance of subsyndromal depression in older primary care patients: prevalence and associated functional disability. *Journal of the American Geriatric Society*, *47*(6), 647–652.

McGivney, S. A., Mulvihill, M., & Taylor, B. (1994). Validating the GDS depression screen in the nursing home. *Journal of the American Geriatric Society*, *42*(5), 490–492.

Mills, T. L., Alea, N. L., & Cheong, J. A. (2004). Differences in the indicators of depressive symptoms among a community sample of African-American and Caucasian older adults. *Community Mental Health Journal*, *40*(4), 309–331.

Rapp, S. R., Parisi, S. A., & Wallace, C. E. (1991). Comorbid psychiatric disorders in elderly medical patients: a 1-year prospective study. *Journal of the American Geriatric Society*, *39*(2), 124–131.

Rosen, J., Mulsant, B. H., & Pollock, B. G. (1997). Depression in the nursing home: risk factors and interventions. *Nursing Home Medicine*, *5*, 156–164.

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist*, *4*, 165–173.

Skarupski, K. A., Mendes de Leon, C. F., Bienias, J. L., Barnes, L. L., Everson-Rose, S. A., Wilson, R. S., et al. (2005). Black-white

differences in depressive symptoms among older adults over time. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(3), P136–142.

Stiles, P. G., & McGarrahan, J. F. (1998). The Geriatric Depression Scale: a comprehensive review. *Journal of Clinical Geropsychology*, *4*, 89–110.

Wall, J. R., Lichtenberg, P. A., MacNeill, S. E., Walsh, P., & Deshpande, S. A. (1999). Depression detection in geriatric rehabilitation: Geriatric Depression Scale short form vs. long form. *Clinical Gerontologist*, *20*(3), 13–21.

Wancata, J., Alexandrowicz, R., Marquart, B., Weiss, M., & Friedrich, F. (2006). The criterion validity of the Geriatric Depression Scale: A systematic review. *Acta Psychiatrica Scandinavica*, *114* (6), 398–410.

Yesavage, J. A. (1988). Geriatric Depression Scale. *Psychopharmacology Bulletin*, *24*(4), 709–711.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatry Research*, *17*(1), 37–49.