Published in final edited form as:

Am J Geriatr Psychiatry. 2008 November ; 16(11): 914–921. doi:10.1097/JGP.0b013e318186bd67.

Screening Performance of the Geriatric Depression Scale (GDS-15) in a Diverse Elderly Home Care Population

Linda G. Marc, ScD, MPH, MS, Patrick J. Raue, PhD, and Martha L. Bruce, PhD, MPH

From the Department of Psychiatry, Weill Medical College of Cornell University, Westchester, NY;
and the Center for Multicultural Mental Health Research, Cambridge Health Alliance-Harvard
Medical School, Somerville, MA

Keywords

depression; differential item functioning; home care; optimal cutoff value; sensitivity; specificity

Introduction

The assessment of depression in elderly homecare patients is essential for determining the magnitude and nature of depression (1); however in clinical practice where time is at a premium, diagnostic instruments like the Structured Clinical Interview for DSM-IV (SCID) (2) are not routinely used in the homecare setting. While screening for depression is part of the comprehensive assessment of homecare patients, there is no information on the validity of standardized screens relative to diagnostic assessment in such populations. This study examines the sensitivity and specificity of the 15-item Geriatric Depression Scale (GDS-15) (3) compared to the SCID, a gold standard assessment.

More than 20 years ago the 30-item GDS was developed as a self-report instrument to screen for clinical depression among the elderly (4). The instrument excludes certain somatic symptoms which might be due to medical illness, and makes use of a simple response format (yes/no, rated 1 or 0) which facilitates easier use by individuals with impaired cognitive functions. The endorsed items are then totaled, generating a score from which patients are classified as depressed or non-depressed. The development, validation and factor structure of the shorter GDS-15 has been described previously, elsewhere (3), and has been evaluated in a variety of inpatient, outpatient, primary care, and nursing home populations (5). While the short form is more practical for use amongst the elderly, its administration to homecare patients burdened with poor medical and functional status has not been reported on. Furthermore, its validity, reliability, sensitivity and specificity, compared to a gold standard have not yet been examined in homebound patients.

Although the use of the GDS score assumes unidimensionality (a single underlying construct of depressive symptoms) and no item-level bias, the effect of independent factors (i.e., age, educational attainment, gender, and race/ethnicity) on the measurement properties of the GDS in *homebound* patients is unknown. While there have been reports that the instrument performs poorly in the “old-old” (6), and amongst persons with low or no formal level of education (7, 8), Tang et al. (9) found no differential item functioning (DIF) across age or education in an elderly population. To our knowledge, there have been no reports on item bias due to the effect

of gender or race/ethnicity on the measurement properties of the GDS. To further examine the effects of these variables on the properties of the GDS-15 in a homebound population we will employ DIF analysis to examine the degree to which items that comprise the scale are systematically related to these independent factors (10,11). As an example, item difficulty bias can be determined across gender if we investigate whether women, compared to men, more frequently respond higher on certain items, after matching the subgroups on level of depression (usually the total scale score) (12). Item discrimination bias is determined by evaluating whether the item difficulty bias increases or decreases as a function of the level of depressive symptoms (the underlying construct). Drawing a parallel from the field of epidemiology to the field of psychometrics, the evaluation of item difficulty (uniform DIF) and item discrimination (non-uniform DIF) are analogous to confounding and effect modification, respectively (13).

Thus, the primary aim of this study was to empirically evaluate the psychometric properties of the GDS-15 in an elderly home health care population, determine the optimal cutoff points and screening performance for the detection of major depression, and to examine age, level of educational, gender and race/ethnicity on the measurement properties of the scale. As distress associated with medical illness and disability in an elderly homebound population may confound the ability of this tool to correctly detect or recognize depression, we hypothesize that the sensitivity and specificity of this instrument will be influenced by severity of medical burden, impaired functioning and cognitive impairment. Finally, as there have been no previous reports to date on item bias analyses in the GDS-15 administered to elderly homebound patients, herein, we propose exploratory DIF analyses.

Methods

The study received full review and approval from the Institutional Review Board of Weill Medical College of Cornell University. All study participants were provided an informed consent for signature.

Participants

This was a secondary analysis of data collected from a random sample study including five hundred and twenty-six subjects aged 65 and older, newly admitted over a 2 year period (Dec. 1997 to Dec. 1999) to a large visiting nurse service agency in Westchester, New York. As the validity of the 15-item GDS in cognitively impaired elderly subjects has been questioned, study patients who scored < 18 on the MMSE were excluded from the psychometric analyses (14).

The sampling strategy to recruit a representative sample of agency patients has been previously described, elsewhere (15). The original study including 539 patients was designed to report on the distribution, course, and outcomes of DSM-IV major depression in elderly patients receiving home care for medical and surgical problems. Selected participants were interviewed by bachelor and master's-level research associates two weeks after admission.

Study subjects had a mean age of 78.3 (SD $\pm$ 7.5) years, and were predominately female (N=351/539; 65.1%). Overall demographic characteristics were diverse across ethnicity, education, marital and poverty status, and very similar to reported national statistics of home care patients (16). Ethnic composition included non-Hispanic White (N=458/539; 85.0%), non-Hispanic Black (N=56/539; 10.4%) and Hispanics (N=21/539; 3.9%). At time of interview, one-third were married (N=204/539; 37.9%); one-fourth (N=94/363) reported living in poverty, as defined by the 1998 guidelines of the U.S. Department of Health and Human Services (17); and there was variation reported across educational attainment: less than high school (N=164; 30.6%); high school (N=170; 31.7%); some college (N=91; 17.0%); college (N=53; 9.9%); post-college (N=58; 10.8%).

Measures

Self-report measures of depression were taken by research assistants using the 15-item GDS. Separate patient and informant interviews assessed current DSM-IV criteria for major depression using the mood module of the SCID. The diagnosis of depression was established by consensus of the study geriatric psychiatrist, geriatrician, clinical psychologist (PJR), and principal investigator (MLB) using clinical information based on all sources of information, including patient interview, informant interview, and patient medical status and medications as documented in the medical record (Health Care Financing Administration form 485). This consensus procedure has been previously described elsewhere, and used in the parent study (15,18,19). An inclusive approach to assigning diagnoses was used, whereby symptoms were rated as present regardless of whether they were due to general medical conditions or medications (20).

Cognitive impairment was assessed using the Mini-Mental State Examination (MMSE) (21), an instrument shown to have consistency and reliability in detecting cognitive functioning in an elderly population (22). Patients were grouped into a dichotomous category indicating cognitive impairment (MMSE score ranging 18–23) or no cognitive impairment (MMSE score ≥ 24), which is the most widely used and accepted cutoff for the MMSE (23). Medical morbidity was determined from the medical record and patient interview by a geriatric internist using the Charlson Comorbidity Index (24). The Charlson is the most extensively studied comorbidity index, which shows strong evidence of moderate-to-high psychometric properties in the disabled and elderly populations (25,26).

Other self-report measures include counts of activities of daily living (ADLs) and instrumental activities of daily living (IADLs) the patient was unable to do without assistance (27). Evidence from the literature suggests that the criterion validity of these indexes are satisfactory when assessed in terms of the correlation with an outcome variable (i.e., home help) (28). Finally, pain intensity was assessed by the single three-level item (“a great deal,” “a little bit,” or “none”) from the Medical Outcomes Study 36-item Short-Form Health Survey (29).

Statistical Analysis

A receiver operating characteristic (ROC) curve was plotted for the GDS-15 and SCID diagnosis to compare the sensitivity and specificity of each threshold for major depression (30). Sensitivity was defined as the probability of a positive screening for depression given that the individual met criteria for depression using information from the SCID. Specificity was defined as a negative screen for depression, given that the individual did not meet the clinical criteria for depression. ROC curves were plotted separately for cognitive impairment, disability and pain. The goal was to compare the sensitivity and specificity of each threshold of morbidity. Threshold categories were based on the median score for each measure (MMSE M=3; Charlson M=2; ADL M=1; IADL M=4; Pain M=2), assigning patients “worse” or “better” health status. Accuracy was measured by the area under the ROC curve (AUC), and a p-value was reported for the statistical test comparing the equality of ROC curves, detecting the difference between areas under the curves for medical conditions and for sociodemographic characteristics.

Optimal cutoff scores were determined using Youden’s Index to summarize the information into a single numeric value (31). Internal consistency was evaluated using the Kuder-Richardson formula 20 (KR-20), a special version of alpha for items that are dichotomous (32). Two-tailed t-tests were used for continuous variables to compare the mean scale score between different subgroups. Chi-square analyses were performed on dichotomous categorical variables. Alpha-level 0.05 was used for determining significance.

DIF was evaluated using logistic regression to predict item responses across dichotomous categories for gender (0=male; 1=female) and nonwhite race (0=white; 1=nonwhite). Ordinal logistic regression, using the proportional odds model was used to predict item responses across the ordinal categorical variables for educational attainment (0=< HS; 1=HS; 2=Some College; 3=College; 4=post-college) and age, which was categorized into a three-level ordinal variable for this analysis (0=65–74 years; 1=75–84 years; 2=85+) (33,34). Measures of magnitude consisted of the odds ratio with 95% confidence intervals. Evidence of DIF was defined as an odd's ratios ≥ 2.0 or conversely ≤ 0.50 (10,12). Items with no evidence of DIF were totaled and scored to assess their overall performance, compared to the original score using the 15-items. Software programs used to manage and analyze the data described herein include STATA (35), SPSS (36) and EXCEL (37).

Results

Examining the study population (excluding cases with moderate to severe dementia, N=13), prevalence of depression using the SCID was 15.4% (N=81/526). A GDS-15 scale was then generated using casewise deletion, dropping 30 subjects who did not respond to all 15 items, and 4 subjects who had not responded to any items. Examining the remaining cases (N=492), bivariate analyses showed that major depression was not significantly associated with any sociodemographic factors, and these findings are similar to results previously reported in the parent study (15).

Optimal Cutoff, Sensitivity and Specificity using the SCID

Preliminary psychometric analyses of the scale, based on 492 study participants, shows an overall mean of 3.5 (ranging from 0 to 13) with an internal consistency-reliability equal to 0.80 (Table 1). Individuals with depression, compared to non-depressed, had significantly higher scores on the GDS-15 ($t=10.23$, $df=490$, $P<0.001$), Charlson Medical Comorbidity Index ($t=2.21$, $df=524$, $p=0.027$), instrumental activities of daily living ($t=2.73$, $df=513$, $p=0.007$) and pain intensity ($t=4.23$, $df=512$, $P<0.001$) (Table 1). In an unadjusted logistic regression model, analyses showed that patients reporting at least three ailments (cluster), compared to patients reporting less or none, were 2.5 times more likely to be diagnosed with major depression (defined by SCID) (Odds Ratio = 2.47; 95% CI 1.49–4.09). These results remained unchanged in an adjusted stepwise logistic regression model (Odds Ratio = 2.47; 95% CI 1.49–4.09; Wald $\chi^2=12.26$; 1df; $n = 524$ with data on all variables) (Table 1).

The optimal cutoff using the trade-off between true-positives and true-negatives (Youden's Index) is 5, with a sensitivity of 71.8%, specificity of 78.2% and ROC area under the curve at 0.793 (Table 2). When the sensitivity and specificity of the GDS-15 were examined across clinical, functional, cluster and demographic variables, using the cutoff of 5, there were no statistical differences across any of these factors. Sensitivity appeared somewhat higher among patients with cognitive impairment, greater medical morbidity, disability and pain but, again, these differences did not reach statistical significance. We also explored whether sensitivity differed by a composite variable indicating patients who had at least three of these conditions, but the difference did not reach statistical significance.

Differential Item Functioning

Results of the DIF analyses indicated that age, educational attainment, gender and race do not influence the measurement properties of the GDS-15. Of the 15 items evaluated, 12 showed no evidence of either uniform or non-uniform DIF and were relatively free of item bias. However, three items met the criteria for evidence of bias, with odd's ratios ≥ 2.0 or conversely ≤ 0.50 (Table 3). Briefly, only uniform DIF was observed for items 5, 10, and 14 by sex; and for item-10 by race.

Results show that the likelihood of women responding negatively to the item pertaining to “Are you in good spirits most of the time” was 3.43 times (OR=3.43; Wald $\chi^2=85.1$; df=2; $p=0.04$) that of men, controlling for total GDS score. In contrast, men were twice as likely than women to respond positively to items – “Do you feel you have more problems with memory” (OR=0.43; Wald $\chi^2=51.2$; df=2; $p=0.04$); and “Do you feel that your situation is hopeless” (OR=0.35; Wald $\chi^2=85.5$; df=2, $p=0.01$), again controlling for total GDS score. Finally, the likelihood of non-whites responding positively to the item, “Do you feel you have more problems with memory” was four times (OR=4.37, Wald $\chi^2=51.0$; df=2; $p<0.001$) that of whites with equivalent GDS scores.

Based on this information above, an analysis of the shortened GDS score was performed deleting the three items with evidence of bias (i.e., items 5, 10 and 14). The psychometric properties of this 12-item version (AUC=0.8681) was not significantly improved over the original 15-item scale (AUC=0.8716). In another attempt to further examine if improvements could be made, a 14 item version of the scale was created omitting item 10 as it showed item-level bias by both gender and race. Properties of this 14-item version showed no significant improvement over the original GDS-15, although the AUC was slightly higher (0.8732). Overall, DIF analyses suggest that age, level of education, gender and race do not have an effect on the measurement properties of the GDS-15 in an elderly homebound population.

Discussion

In this study population of elderly home health care patients, the psychometric properties of the GDS-15 were similar to other published reports on its sensitivity, specificity and optimal cutoff value. The analyses confirm that using the cutoff of 5 yields optimal sensitivity 71.8%, and specificity 78.2%, when compared to SCID criteria for depression. Hence, the findings in this patient population are similar to those reported in pooled studies of the GDS-15 indicating a sensitivity 80.5%, specificity 75.0%, with optimal cut-off values 5/6 (5).

Although we found some evidence that the GDS-15 (using cutoff of 5) has higher sensitivity and lower specificity in patients with greater comorbidity, these findings were not statistically significant. We also found no evidence that the accuracy of the GDS-15 was influenced by sociodemographic factors. Analyses were performed to compare areas under the curve across the diverse subgroups of this population, showing the instrument performs comparably, increasing the generalizability of these study results. In addition, differential item functioning analyses revealed no variability of item responses across subgroups identified by age, level of education, gender or race. Thus, these results suggest that the GDS-15 offer no evidence of a difference across the middle aged, elderly and “old-old” (≥ 75 years), across low versus high levels of educational attainment, nor across gender or race.

We did find some evidence of sociodemographic variation in which symptoms were endorsed after matching for total GDS score. Women were less likely than men to endorse being in “good spirits most of the time”. Again matching for total GDS, men were more likely than women, and non-whites were more likely than whites, to endorse memory problems. But omitting these items from the total GDS score did not improve the score’s psychometric properties. Whether or not these differences in item endorsement across demographic groups is clinically meaningful is another question that could be pursued with different types of analyses in the future.

Limitations

A potential limitation to the generalizability of these findings is that the study sample was drawn from a single visiting nurse service agency in Westchester, New York. However, the sociodemographic and clinical characteristics of the study sample reflect those of home

healthcare patients nationally, suggesting that the findings have broad relevance. A second limitation is the high rate of nonparticipation among sample patients that, while common to studies of homebound seniors, results in potential selection bias. We do not know if these findings apply to patients who did not participate. A third limitation is that patients were grouped into cognitive impairment vs. no cognitive impairment, using an MMSE cut-off score of 23/24 (23); however without more in-depth clinical evaluation, some subjects labeled "no cognitive impairment" may in fact have had "mild cognitive impairment" of various nosologies.

Conclusions

The Geriatric Depression Scale was developed to give a simple, easy to use approach to screening for depression in older adults. The advantage of the GDS for medically ill populations is that the instrument purposely does not assess the somatic symptoms of depression as to not inflate the total score by inadvertently attributing symptoms of medical illness to depression. A risk in this approach is that the scale might underestimate cases of depression by systematically excluding those symptoms of depression that are somatic. Our data suggest, however, that both the sensitivity and specificity of the GDS are well within acceptable ranges. Further, accuracy of the GDS-15 is not influenced by severity of medical burden, age or other sociodemographic characteristics even in a medically ill and disabled patient population.

These results have broad implications for depression screening suggesting that (i) the "very old" and ill can be screened appropriately despite clinician beliefs that this population is too difficult to assess; (ii) the presence of a major depressive episode among elderly homebound adults can be reliably detected; and (iii) the tool is useful for the detection of depression across culture-specific populations (38,39).

Acknowledgements

Disclosures: LGM (#T32 MH19132 and #T32 MH067555), PJR (#K23 MH069784) and MLB (#R01 MH 56482) were supported, in part, by grants from the National Institute of Mental Health. Additional support for methodological consultation was provided to LGM as a 2006-7 Scholar of the African-American Mental Health Research Scientist Consortium (AAMHRS), and 2006 Program Scholar in the Summer Institute for Applied Multi-Ethnic Research, at the Inter-University Consortium for Political and Social Science Research, University of Michigan, Ann-Arbor. LGM received consultancy fees from Behavioral Science International LLC, GlaxoSmithKline and Pfizer during the study period.

References

1. Lyness JM, Noel TK, Cox C, King DA, Conwell Y, Caine ED. Screening for depression in elderly primary care patients. A comparison of the Center for Epidemiologic Studies-Depression Scale and the Geriatric Depression Scale. *Arch Intern Med* 1997 Feb 24;157(4):449-454. [PubMed: 9046897]
2. Spitzer, RL. User's guide for the Structured clinical interview for DSM-III-R : SCID. Washington, DC: American Psychiatric Press; 1990.
3. Sheikh, JI.; Yesavage, JA. Geriatric Depression Scale (GDS). Recent evidence and development of a shorter version. In: Brink, TL., editor. *Clinical Gerontology : A Guide to Assessment and Intervention*. New York: The Haworth Press; 1986. p. 165-173.
4. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982;17(1):37-49. [PubMed: 7183759]
5. Wancata J, Alexandrowicz R, Marquart B, Weiss M, Friedrich F. The criterion validity of the Geriatric Depression Scale: a systematic review. *Acta Psychiatr Scand* 2006 Dec.;114(6):398-410. [PubMed: 17087788]
6. Watson LC, Lewis CL, Kistler CE, Amick HR, Boustani M. Can we trust depression screening instruments in healthy 'old-old' adults? *Int J Geriatr Psychiatry* 2004 Mar.;19(3):278-285. [PubMed: 15027044]

7. Cwikel J, Ritchie K. Screening for depression among the elderly in Israel: an assessment of the Short Geriatric Depression Scale (S-GDS). *Isr J Med Sci* 1989 Mar.;25(3):131–137. [PubMed: 2708013]
8. Kim JM, Prince MJ, Shin IS, Yoon JS. Validity of Korean Form of Geriatric Depression Scale (KGDS) among cognitively impaired Korean elderly and the development of a 15-item short version (KGDS-15). *Int J Methods Psychiatr Res* 2001;10:204–210.
9. Tang WK, Wong E, Chiu HF, Lum CM, Ungvari GS. The Geriatric Depression Scale should be shortened: results of Rasch analysis. *Int J Geriatr Psychiatry* 2005 Aug;20(8):783–789. [PubMed: 16035120]
10. Cole SR. Assessment of differential item functioning in the Perceived Stress Scale-10. *Journal of Epidemiology & Community Health* 1999;53(5):319–320. [PubMed: 10396541]
11. Holland, PW.; Wainer, H. Differential Item Functioning. Hillsdale, NJ: Lawrence Erlbaum; 1993.
12. Cole SR, Kawachi I, Maller SJ, Berkman LF. Test of item-response bias in the CES-D scale. experience from the New Haven EPESE study. *Journal of Clinical Epidemiology* 2000;53(3):285–289. [PubMed: 10760639]
13. Crane, P. (Personal Communication) DIFdetect for STATA. Measurement Design and Statistical Methods for Health Outcomes Research Seminar. Boston, MA: Harvard School of Public Health; 2003 Feb 14. An extension of ordinal logistic regression for DIF detection.
14. Feher EP, Larrabee GJ, Crook TH 3rd. Factors attenuating the validity of the Geriatric Depression Scale in a dementia population. *J Am Geriatr Soc* 1992 Sep;40(9):906–909. [PubMed: 1512387]
15. Bruce ML, McAvay GJ, Raue PJ, Brown EL, Meyers BS, Keohane DJ, et al. Major depression in elderly home health care patients. *Am J Psychiatry* 2002 Aug;159(8):1367–1374. [PubMed: 12153830]
16. Haupt BJ, Jones A. The National Home and Hospice Care Survey: 1996 summary. *Vital Health Stat* 13 1999 Oct;141:1–238. [PubMed: 10549237]
17. Federal Register 1998 February 24;:9235–9238. [PubMed: 10177502]
18. Leckman JF, Sholomskas D, Thompson WD, Belanger A, Weissman MM. Best estimate of lifetime psychiatric diagnosis: a methodological study. *Arch Gen Psychiatry* 1982 Aug;39(8):879–883. [PubMed: 7103676]
19. Klein DN, Ouimette PC, Kelly HS, Ferro T, Riso LP. Test-retest reliability of team consensus best-estimate diagnoses of axis I and II disorders in a family study. *Am J Psychiatry* 1994 Jul;151(7):1043–1047. [PubMed: 8010362]
20. Koenig HG, George LK, Peterson BL, Pieper CF. Depression in medically ill hospitalized older adults: prevalence, characteristics, and course of symptoms according to six diagnostic schemes. *Am J Psychiatry* 1997 Oct;154(10):1376–1383. [PubMed: 9326819]
21. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975 Nov;12(3):189–198. [PubMed: 1202204]
22. Lopez MN, Charter RA, Mostafavi B, Nibut LP, Smith WE. Psychometric properties of the Folstein Mini-Mental State Examination. *Assessment* 2005 Jun;12(2):137–144. [PubMed: 15914716]
23. Folstein, MF.; Folstein, SE.; McHugh, PR.; Fanjiang, G. Mini-Mental State Examination user's guide. Odessa: Florida; 2001.
24. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40(5):373–383. [PubMed: 3558716]
25. de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity. a critical review of available methods. *J Clin Epidemiol* 2003 Mar;56(3):221–229. [PubMed: 12725876]
26. Charlson ME, Peterson JC, Syat BL, Briggs WM, Kline R, Dodd M, et al. Outcomes of community-based social service interventions in homebound elders. *Int J Geriatr Psychiatry*. 2007 Oct 4;
27. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969 Autumn;9(3):179–186. [PubMed: 5349366]
28. Norstrom T, Thorslund M. The structure of IADL and ADL measures: some findings from a Swedish study. *Age Ageing* 1991 Jan;20(1):23–28. [PubMed: 2028846]

29. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992 Jun;30(6):473–483. [PubMed: 1593914]
30. Erdreich LS, Lee ET. Use of relative operating characteristic analysis in epidemiology. A method for dealing with subjective judgement. *Am J Epidemiol* 1981 Nov;114(5):649–662. [PubMed: 7304595]
31. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32–35. [PubMed: 15405679]
32. Kuder GF, Richardson MW. The Theory of the estimation of test reliability. *Psychometrika* 1937;2:151–160.
33. Ananth C, Kleinbaum D. Regression models for ordinal responses: a review of methods and applications. *Int J Epidemiol* 1997;26(6):1323–1333. [PubMed: 9447413]
34. Swaminathan H, Rogers H. Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement* 1990 Win;27(4):361–370.
35. Stata [computer program] Release. Vol. 9. ed. College Station, Tex.: Stata Corp.; 2005.
36. SPSS for Windows 14.0 [computer program]. Chicago, Ill: SPSS Inc.; 2005.
37. Microsoft Excel [computer program]. Microsoft Corp.; 2000.
38. Rait G, Burns A, Baldwin R, Morley M, Chew-Graham C, St Leger AS, et al. Screening for depression in African-Caribbean elders. *Fam Pract* 1999 Dec.;16(6):591–595. [PubMed: 10625132]
39. Harralson TL, White TM, Regenberg AC, Kallan MJ, Have TT, Parmelee PA, et al. Similarities and differences in depression among black and white nursing home residents. *American Journal of Geriatric Psychiatry* 2002 Mar–Apr;10(2):175–184. [PubMed: 11925278]

Table 1
Mean Scores of Clinical and Functional Factors by Depression Status Among Elderly Home Health Care Patients

Major Depression [†]							
Factor	Alpha	Yes		No		Analysis	
		Mean	SD	Mean	SD	t	p
Geriatric Depression Scale (GDS-15, range=0–13; N=492)	KR20=0.80	6.7	3.5	3.0	2.6	10.23	<0.001
Medical Morbidity (Charlson Comorbidity Index, range 0–10)		3.3	2.3	2.5	2	2.21	<0.027
Activities of Daily Living Disability – ADLD (range=0–6)		1.3	1.5	1.0	1.2	0.54	0.5913
Instrumental Activities of Daily Living Disability – IADLD (range = (0–6)		3.7	1.4	3.2	1.5	2.73	0.007
Cognitive Function (MMSE, range=0–30)		26.3	3.2	26.6	2.8	–0.76	0.450
Reported Pain (range=1–3)		2.3	0.8	1.9	0.8	4.23	<0.001
Cluster [‡]	Depression		No Depression				
	N	%	N	%			
Subjects reporting cluster including medical morbidity, ADL, IADL or pain		31	38.3	89	20.0	Adj OR=2.47 (95% CI 1.49–4.09) Wald χ^2 =12.26, 1df, p=0.005	
Subjects reporting none or less than three comorbid factors		50	61.7	356	80.0		

[†]Depression defined using SCID

[‡]Subjects reporting at least three of the four comorbid factors (Charlson Index, ADLD, IADL or pain)

Table 3
Differential Item Functioning: Odds Ratios for GDS-15 Items Across Categories for Age, Education, Gender and Race

	Age										Education										Gender										Race									
	Uniform					Non-Uniform					Uniform					Non-Uniform					Uniform					Non-Uniform					Uniform					Non-Uniform				
	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p
Q1	0.80	100.7	2	0.26	0.78*	93.1	3	<0.001	1.17	104.4	2	0.14	1.12*	111.7	3	0.03	1.08	87.7	3	0.30	1.05	104.8	2	0.07	0.99	113.3	3	0.88	1.04	113.3	3	0.27	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q2	1.06	82.7	2	0.68	0.97	84.9	3	0.80	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	105.5	2	0.90	1.01	105.5	3	0.83	0.79	113.3	3	0.88	1.04	113.3	3	0.27	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q3	1.03	105.5	2	0.90	1.01	105.5	3	0.83	0.79	113.3	3	0.87	1.04	113.3	3	0.27	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q4	0.89	108.2	2	0.46	0.99	113.7	3	0.87	1.04	113.3	3	0.27	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53
Q5	0.97	91.5	2	0.92	0.97	90.9	3	0.65	1.04	91.3	2	0.84	1.05	91.1	3	0.33	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q6	0.98	75.8	2	0.93	1.09	76.6	3	0.10	1.16	78.3	2	0.20	0.98	79.4	3	0.51	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q7	0.76	84.8	2	0.21	0.97	87.4	3	0.65	1.44*	78.6	2	0.01	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53
Q8	1.29	102.1	2	0.13	1.14	106.5	3	0.11	0.84	102.9	2	0.07	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30
Q9	1.22	59.4	2	0.16	0.96	61.4	3	0.48	0.91	61.4	2	0.25	1.05	61.7	3	0.11	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q10	1.02	51.8	2	0.94	1.00	51.9	3	0.98	0.97	51.5	2	0.85	1.00	52.5	3	0.78	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q11	1.72	67.1	2	0.09	0.92	67.3	3	0.34	1.31	71.4	2	0.13	0.95	70.3	3	0.34	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q12	0.82	116	2	0.35	1.14	118.5	3	0.17	0.79	126.7	2	0.08	0.89*	130.6	3	0.01	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q13	0.87	50.6	2	0.31	0.99	51.4	3	0.87	1.28*	58.7	2	0.002	1.09	54.2	3	0.17	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q14	1.12	82.6	2	0.65	1.07	80.6	3	0.36	0.85	80.9	2	0.28	1.05	82.6	3	0.22	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75
Q15	1.19	95.4	2	0.37	0.95	97.6	3	0.42	0.70*	97.9	2	0.01	1.00	98.3	3	0.82	1.03	79.6	3	0.69	1.01	103.2	3	0.75	1.05	82.3	2	0.53	1.08	87.7	3	0.30	1.03	79.6	3	0.69	1.01	103.2	3	0.75

	Age						Education					
	Uniform			Non-Uniform			Uniform			Non-Uniform		
	OR	χ^2	df	p	OR	χ^2	df	p	OR	χ^2	df	p
Q5	3.43 [†]	85.1	2	0.04	1.15	85.3	3	0.30	1.21	91.6	2	0.78
Q6	1.31	77.4	2	0.41	0.86	81.5	3	0.11	0.61	77.0	2	0.29
Q7	1.33	79.5	2	0.45	1.21	76.2	3	0.19	0.66	79.9	2	0.38
Q8	1.67	103.7	2	0.08	1.21	105.3	3	0.11	0.65	104.5	2	0.34
Q9	1.06	59.1	2	0.78	0.95	63.1	3	0.50	1.30	59.4	2	0.37
Q10	0.43 [†]	51.2	2	0.04	0.86	50.1	3	0.13	4.37 [†]	51.0	2	<0.001
Q11	1.37	67.3	2	0.48	0.68 [*]	80.5	3	0.01	1.07	71.0	2	0.91
Q12	0.84	119.9	2	0.59	1.00	120.3	3	1.00	1.41	119.3	2	0.47
Q3	1.12	51.5	2	0.60	0.98	52.3	3	0.88	0.78	51.3	2	0.37
Q14	0.35 [†]	85.5	2	0.01	0.99	88.2	3	0.94	1.00	82.1	2	1.00
Q15	0.53 [*]	94.4	2	0.04	0.80	92.4	3	0.06	1.52	95.6	2	0.32

* Statistically significant, does not meet DIF criteria

[†] Statistically significant, meets DIF criteria

OR=Odds Ratio; χ^2 =Wald Statistic

df=degrees of freedom; p= p-value

Demographic Categories:

Age (0=65–74 years; 1=75–84 years; 2=85+);

Education (0=< HS; 1=HS; 2=Some College; 3=College; 4=Post-College)

Gender (0=male; 1=female)

Race (0=white; 1=nonwhite)