INTERNATIONAL JOURNAL OF GERIATRIC PSYCHIATRY
Int J Geriatr Psychiatry 2006; **21**: 588–593.
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/gps.1533

Reliability, validity and factor structure of the GDS-15 in Iranian elderly

Seyed Kazem Malakouti[1]*, Paridokht Fatollahi[2], Arash Mirabzadeh[3], Mojgan Salavati[3] and Taher Zandi[4]

[1]*Department of Psychiatry, Iran University of Medical Sciences*
[2]*Ministry of Oil, Health Department, Iran*
[3]*Department of Psychiatry, University of Welfare and Social Sciences, Iran*
[4]*State University of New York, Distinguished Professor, Director of Mood and Memory Clinic*

SUMMARY

Objectives Depressive disorders are a public health problem even in developing countries. Access to valid and reliable screening instruments is needed for conducting community surveys. The main objective of this study is to provide the Iranian version of the Geriatric Depression Scale-15 (GDS).

Methods The GDS-15 Farsi version was developed by translation and back translation. Two hundred and four subjects aged 59 years or older, who were chosen randomly from residents of the Ekbatan district of Tehran, the capital city of Iran, completed the GDS-15. The Composite International Diagnostic Interview (CIDI) was used to establish a gold standard diagnosis of major depressive disorders.

Results The GDS was found to be an internally consistent measure. Alpha, split-half coefficients and test–retest reliability were 0.9, 0.89 and 0.58 respectively. Two factors were extracted by using factor analysis and the principle component analysis (varimax rotation): 'depression' and 'psychosocial activity'. The Depression factor (omitting items 2, 9, 10, 13), which could be considered as a short form of the scale ($\alpha = 0.92$), has significant correlation with the main scale ($r = 0.58$). Using receiver operating curve (ROC) analysis, the optimum cutoff score for GDS-15 is 7/8, yielding a sensitivity of 0.9 and a specificity of 0.84. The optimum cutoff score for GDS-11 is 6, yielding a sensitivity of 0.9 and a specificity of 0.83.

Conclusion The long and short forms of the GDS have excellent properties as screening instruments for major depression in older dwellers in Iran, particularly in urban areas, as presented in our findings. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS — depression; elderly; screening; geriatric depression scale-15

INTRODUCTION

The world population is growing old. Today, about two-thirds of all older people are living in developing countries (WHO, CHP). According to the latest Iranian census data (1996), 3.2% and 1% of the population are between 65–74 and older than 75 and 6.5% are older than 60 respectively and it is estimated that by 2020 the elderly population will be about 10% (Iran Ministry of Health, Family Health Office, 1999). Epidemiological studies have revealed that depressive syndromes are most common among elderly in different settings (Reynolds, 1996; Lebowitz *et al*., 1997; Steffens *et al*., 2000; Alexopoulous, 2001).

Different instruments have been used for screening depression in older people (Gareri *et al*., 2002). Among them, the Geriatric Depression Scale (GDS), developed by Yesavage and Brink (1983), is used extensively.

The GDS-15 has been translated into many languages and validated in Brazil (Almeida and Almeida, 1999), China (Boey, 2000), UK (Arthur

*Correspondence to: Dr S. K. Malakouti, 82, Mahyar st. Vali-e-Asr Ave, Tehran 19669 Iran. Tel/fax: (0098-21) 8718914-8706306.
E-mail: kmalakouti@iums.ac.ir or malakoutik@yahoo.com

Contract/grant sponsor: Mental Health Research Centre of Tehran Psychiatric Institute; contract/grant number: 22.

Received 7 July 2005
Accepted 16 November 2005

et al., 1999), Malay (The EE, online paper), Netherlands (De Craen *et al*., 2003).

Using Receiver Operating Curve (ROC) analysis, and gold standards such as the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID), Composite International Diagnostic Interview (CIDI) and International Classification of Disease (ICD-10) research criteria, optimum cutoff scores of 5, 6 and 7 for GDS-15, with Area Under Curve (AUC) between 0.76 to 0.935 and sensitivity with a range of 73–97% and specificity with a range of 60–96% were obtained from studies (Lyness *et al*., 1997; Pomeroy *et al*., 2001; Robison *et al*., 2002; Schreiner *et al*., 2003). The samples for these studies were recruited from primary care settings or rehabilitation units.

In this study, the GDS-15 version has been selected in order to develop an Iranian version based on a community sample.

METHODS

Materials

For this study, the Ekbatan district of Tehran was selected due to its higher than average number of elderly residents. Under supervision of the directors of every residential block and the Health Centre, the Health volunteers gathered the census by door-to-door survey.

Participants

From 1422 old people, 300 subjects were chosen randomly. During the six months between taking the census and completing the questionnaires, 54 subjects refused to participate due either to death or moving to a new location, and 42 subjects refused to participate in this research. Finally, 204 subjects were enrolled and completed the GDS-15. From the study sample, 104 subjects were selected randomly and interviewed by CIDI to determine the cutoff score. No significant difference was seen in sex between enrolled and omitted groups. The age of 13.5% ($n = 13$) of omitted samples was more than 85 where as, in the study sample it was 3.9% ($n = 8$) ($\chi^2 = 9.6\%$, df $= 2$, $p < 0.01$).

By telephone call, the project was described briefly to the subjects and after obtaining their consent; appointments were made at their home.

Instruments

The GDS-15 short form was developed from the GDS-30 (Yesavage and Brink, 1983). Its validity was confirmed by other studies in China, UK and Malay with a range of 0.7–0.9 (Lee *et al*., 1993; D'Ath *et al*., 1994; Mui, 1996; Boey, 2000; The EE, online paper). Test–retest reliability (range: 0.7–0.84) and split-half reliability (0.82) were acceptable.

The Farsi version of GDS-15 was developed through translation and back translation process. Twenty-five subjects of 60 or higher in the 13th district of Tehran completed the first version in a joint meeting, which was translated by two specialists into Farsi. Slight modification in wording of items 9, 11, 12 and 13 were required. Twenty-six elderly in Ekbatan district completed the second version again. The final version was translated into English and compared with the original one. In a joint session the final version was developed with two English language specialists and a psychiatrist and a psychologist.

The CIDI has proven to have excellent validity and reliability in 20 countries (WHO, 1990; Andrews and Peters, 1998) and satisfactory diagnostic agreement with clinical diagnosis (Janca *et al*., 1992).

In a study in Iran, the validity indices of CIDI in terms of agreement with diagnosis made by the clinician using diagnostic checklists, had a sensitivity of 0.91 and specificity of 0.53 with positive, negative likelihood ratio of 1.94 and 0.17 respectively. The reliability indices for test–retest evaluation (kappa $= 0.55$, Yule's $Y = 0.56$) were acceptable (Ahmadi *et al*., 2003).

After making an appointment, the interviewers, including two psychiatrists and one psychologist (MS), administered the instruments. Two hundred and four subjects, of whom 104 were interviewed by CIDI, completed the GDS. Given that some study subjects were illiterate, the interviewers completed the GDS by reading the items to the interviewees. Twenty-eight subjects were selected for test–retest reliability measurement after 2 weeks. Using CIDI lifetime version in this study, for providing the number of subjects who were suffering from major depressive disorder (MDD) at the time of completing GDS, only item 1 of E27 question was considered. This item confirms the presence of MDD in the last 2 weeks.

The human subject committee of Tehran Psychiatric Institute approved this project.

Analyses

Descriptive statistics, Cronbach's alpha, split-half and test–retest methods of reliability was conducted. Exploratory factor analysis (Eigenvalues greater than 1) and criterion validity were used for validity study. Factor loading of 0.4 or greater was considered. A diagnosis of MDD, as indicated by the CIDI, was

Table 1. Demographic feature of the sample ($n = 204$)

Variables Groups	Number	%
Gender		
Female	109	53.4
Male	95	46.6
Age		
59–74	128	62.7
75–84	68	33.3
=>85	8	3.9
Living with		
Partner	71	34.8
Partner & Children	64	31.4
Children & Others	43	21.1
Alone	26	12.7
Employment status		
Retired & Pensioner	124	60.7
Unemployed	75	36.8
Employed	5	2.5
Educational status		
Illiterate	29	14.2
Primary	54	26.5
Middle School	49	24
Diploma/Higher	72	35.2

the gold standard. ROC was generated to visualize the sensitivity and specificity of depression scores.

Findings

Demographic features of the sample ($n = 204$) are presented in Table 1.

GDS score in women ($M = 5.13$, $SD = 4.55$) was higher than men ($M = 3.67$, $SD = 4.18$) ($t = 2.38$, $p < 0.05$). Age did not have any significant effect on GDS score, but with Kruskal-Wallis analysis GDS score was higher in lower educational level ($\chi^2 = 15$, $p < 0.01$). Other demographic parameters did not have any significant effect on GDS score.

Reliability

Cronbach's coefficient and inter-items correlation were 0.9 and 0.4 respectively. The alpha coefficient increased to 0.91 if the questions numbered 2, 9, and 10 were omitted from the scale (Table 2). The split-half coefficient was 0.89 and test–retest reliability after 2 weeks was 0.58 ($p < 0.001$). Students *t*-test analysis showed no significant differences between the scores of the two stages.

Validity

KMO coefficient was 0.9 and Barttelet analysis used with a satisfactory result ($\chi^2 = 1537.11$, $p < 0.001$). Using factor analysis, two factors emerged. The first factor (depression), which included 11 items of the GDS, accounted for 49.1% of the variance in the model (Eigenvalue = 1) and the second factor (psychosocial activities), which included three items, accounted for 9% of the variances. Item 10 loaded less than 0.4 (Table 3). Cronbach's alpha for the two factors were 0.92 and 0.52 respectively and the correlation was 0.5 ($p < 0.001$).

For criterion validity GDS mean scores were compared between the depressive group ($M = 11$, $SD = 2.3$) and the non-depressed ($M = 3.93$, $SD = 4.31$, $p < 0.001$). Moreover, GDS score was correlated significantly with a number of MDD ($r = 0.35$) and dysthymia criterions ($r = 0.4$) (DSM-IV-TR criteria) that emerged from CIDI.

Table 2. The mean, SD and item-total correlation of GDS-15 in Ekbatan district

Items	Mean	SD	Item-total correlation	Alpha after omitting the items
1	0/13	0/33	0/59	0/903
2	0/37	0/48	0/29	0/914
3	0/24	0/43	0/68	0/899
4	0/42	0/49	0/65	0/90
5	0/35	0/48	0/74	0/867
6	0/27	0/44	0/7	0/898
7	0/39	0/49	0/73	0/897
8	0/22	0/42	0/72	0/898
9	0/45	0/5	0/38	0/91
10	0/2	0/4	0/33	0/91
11	0/21	0/41	0/7	0/899
12	0/2	0/4	0/66	0/90
13	0/47	0/5	0/55	0/90
14	0/23	0/42	0/8	0/895
15	0/27	0/45	0/57	0/90
Total score	4/45	4/43		

Table 3. Factor structure of the GDS-15 using exploratory factor analysis

No.	Items	Factors	
		First factor (depression)	Second factor (psychosocial activities)
1	Are you basically satisfied with your life?	0.6	
2	Have you dropped many of your activities and interests?		0.77
3	Do you feel that your life is empty?	0.78	
4	Do you often get bored?	0.62	
5	Are you in good spirits most of the time?	0.72	
6	Are you afraid that something bad is going to happen to you?	0.7	
7	Do you feel happy most of the time?	0.69	
8	Do you often feel helpless?	0.76	
9	Do you prefer to stay at home, rather than going out and doing new things?		0.65
11	Do you think it is wonderful to be alive now?	0.83	
12	Do you feel pretty worthless the way you are now?	0.79	
13	Do you feel full of energy?		0.66
14	Do feel that your situation is hopeless?	0.8	
15	Do you think that most people are better off than you are?	0.64	

Item 10 (Do you feel you have more problems with memory than most?) with factor loading of less than 0.4 was not loaded on any factors.

Table 4 shows the results from the ROC analyses. The area under the curve (AUC) values for GDS-15 is 0.887 ($p < 0.0001$). The optimum cutoff score for GDS-15 is 7/8, with a sensitivity of 0.9 and specificity of 0.84. Positive predictive value (PPV) was measured equal to 36% (Table 5).

The depression factor (11 items) could be considered as a short form of the GDS scale, which correlation coefficient with the GDS-15 was 0.9. Its correlation with MDD and dysthymia were 0.37 and 0.42. The optimum cutoff score for GDS-11 is 6, with a sensitivity of 0.9 and specificity of 0.83 (Figure 1).

DISCUSSION

Using different statistical methods to evaluate reliability and validity, the results showed that the GDS-15 has excellent properties as a screening instrument for MDD in older dwellers in Iran. These kinds of instruments are required for conducting national epidemiological study for elderly. On the other hand, it has been revealed that Iranian people who are suffering from mental disorders prefer to see the GPs or other specialists rather than psychiatrists, as a first preference (Shahmohhamadi, 1990). Having an Iranian version of GDS would be helpful in the primary care setting to screen the elderly for depression or treat the depressive states as comorbid conditions.

This study used ROC method and a validated diagnostic gold standard in a sample of older, city dwellers.

The Cronbach's alpha yielded in this study is compatible with Sheikh (1986) and study results from other countries (Lee *et al*., 1993; D'Ath *et al*., 1994; Mui, 1996; Almeida and Almeida, 1999; Boey, 2000; The EE, online paper).

Table 4. Cutoff scores, sensitivity and specificity of GDS-15 against CIDI using ROC analysis

Scales	Cutoff scores	Sensitivity	Specificity
GDS-15	1.5	1	0.36
	2.5	1	0.53
	3.5	1	0.63
	4.5	1	0.7
	5.5	1	0.75
	6.5	1	0.79
	7.5	0.91	0.8
	8.5	0.91	0.83
	9.5	0.818	0.86
	10.5	0.545	0.86
	11.5	0.364	0.89
	12.5	0.273	0.91

Table 5. Positive and negative predictive values of GDS 15 items for Major Depressive Disorder

GDS		MDD		
		Yes	No	Sum
GDS-15 (Cutoff point = 7/8)	Positive No.	10	18	28
	%	35.7	64.3	100
	Negative No.	1	73	74
	%	1.4	98.6	100

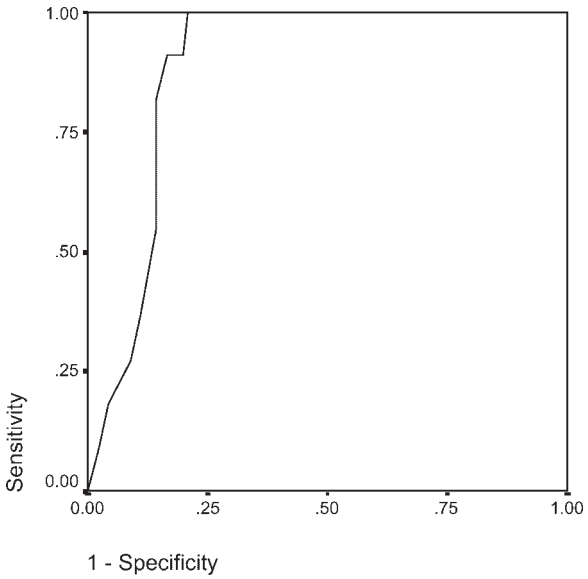

Figure 1. ROC curve of GDS-15 for major depression

Compared with the earlier studies, the results of this study yielded higher cutoff scores (Lyness *et al*., 1997) with better sensitivity and specificity (Robison *et al*., 2002). In the last study, although primary care setting consumers were used as the subjects, lower sensitivity and specificity could be caused by cultural differences in which depressive symptoms might be over-reported by females Puerto Rican who were over-represented in that study.

Given the higher prevalence rate of depression in primary health care settings (Alexopoulous, 2001), using this population as a sample (Lyness *et al*., 1997; Robison *et al*., 2002) could account for lower cutoff point scores. Contrary to this, in other studies, which were performed in the community (Arthur *et al.*, 1996; de Craen *et al*., 2003), the cutoff scores were lower. This could be accounted for by the age range ($>=75$ years) of the subjects, in which the prevalence of depression might be higher. The results of the current study showed a positive relationship between age and GDS scores, although it was not significant. Therefore, high cut-off scores in this study could account for the two sample characteristics: community sample and higher level of education.

The results of this study revealed that the positive predictive value of Iranian versions of GDS-15 is less than 50%. This is not concordant with the results of Almeida and Almeida (1999), which used outpatient subjects. This might be the result of a lower rate of MDD in the community. Lower rate of a disease is accompanied with lower value of PPV (positive predictive value) and NPV (negative predictive value) (Dawson and Trapp, 2000).

Two factors emerged from the factor analysis of GDS-15. The first factor has the most correlation with depression and accounted for nearly 50% of the variance. Considering the items of the first factor, the short-form of GDS-11 was generated with better reliability and validity indices.

It seems that these items (2, 9, 10, 13) might not be implicated in the depression concept in our culture but they should be reevaluated in a larger sample. Social changes in old age, such as retirement and reduction of many daily activities are the most common events in Iranian families. Corresponding results emerged from Malay (The EE, online paper) and China (Mui, 1996), in which items 9 and 13 had no discriminatory value or were changed for cultural adaptation.

Some methodological concerns might affect the interpretation of our data. This study was conducted in one district of the capital, with a relatively small sample. Demographic characteristics of the subjects in this study are different from the old population of Iran in some aspects. According to the latest Iranian general census (1996), 43% are settled in rural areas and more than 70% are illiterate (in this study 14%). In rural areas, the prevalence of depression is lower than urban areas (Rihmer *et al*., 2005) and the results of the current study showed significant reverse relationship between illiteracy and GDS scores. Given the results of this study, the Geriatric Depression Scale (15 items) could be used for screening and clinical studies, particularly in urban areas, but replication studies are recommended using mixed populations and larger sample size.

Another methodological concern is that the CIDI has not been validated in Iranian elderly. According to the CIDI administrative instruction a trained layperson is eligible to conduct the CIDI interview for financial purposes. In this study, using two trained specialists would increase the expenses of the project but not the validity of the CIDI results.

ACKNOWLEDGMENTS

We offer our thanks to Dr. Roohangiz Jamshidi for the CIDI data analysis and excellent review and recommendation. We are grateful to the subjects for their selfless support of this study. Mental Health Research Centre of Tehran Psychiatric Institute supported this work with grant no. 22.

KEY POINTS

- The Iranian version of GDS-15 has excellent reliability and validity as a screening instrument.
- The optimum cutoff score is 7/8 with a high sensitivity and specificity.
- The high cut-off score in this study could be accounted for by the community sample and higher level of subjects' education.
- By omitting four items with low factor loading, GDS-11 was derivated.

REFERENCES

Ahmadi AA, Razaghi O, Sadeghi M, *et al.* 2003. *Mental health effect of Iraqi invasion of Kuwait in war-torn population of Iran*. Ministry of Health and Medical Education. Islamic Republic of Iran.

Alexopoulus G. 2001. *Pharmacotherapy of Depressive Disorders in Older Patients*. McGraw-Hill Health Care information: Minneapolis, MN.

Almeida OP, Almeida SA. 1999. Short version of Geriatric Depression Scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *Int J Geriatr Psychiatry* **14**: 858–865.

Al-Shammari SA, Al-Subaie A. 1999. Prevalence and correlates of depression among Saudi elderly. *Int J Geriatr Psychiatry* **14**: 739–747.

Andrews G, Peters L. 1998. The psychometics properties of the Composite International Diagnostic Interview. *Soc Psychiatry Psychiatr Epidemiol* **33**: 80–88.

Arthur A, Jagger C, Lindesay J. 1999. The 15 item Geriatric Depression Scale had perfect sensitivity but low specificity for detecting depression in elderly people. *Int J Geriatr Psych* **14**: 431–439.

Boey KW. 2000. The use of GDS-15 among the older adults in Beijing. *Clin Gerontol* **21**: 49–60.

D'Ath P, Katona P, Mullan E. 1994. Screening detection and management of depression in elderly primary care attenders. I: the acceptability and performance of the 15-item Geriatric Depression Scale (GDS-15) and the development of short versions. *Fam Pract* **11**: 260–266.

Dawson B, Trapp RG. 2000. *Basic and Clinical Biostatistics*, 3rd edn. Appleton and Lange, USA.

De Craen AJ, Heeren TJ, Gussekloo J. 2003. Accuracy of the 15-item geriatric depression scale (GDS-15) in a community sample of the oldest old. *Int J Geriatr Psychiatry* **18**: 63–66.

Gareri P, De Fazio P, De Sarro G. 2002. Neuropharmachology of depression in aging and age-related diseases. *Aging Res Rev* **1**: 113–134.

Iran Ministry of Health and Medical Education, Family Health Office. 2000. Health of old people in Islamic Republic of Iran.

Janca A, Robins LN, Bucholz KK. 1992. Comparison of Composite International Diagnostic Interview and clinical DSM-III-R criteria checklist diagnosis. *Acta Psychiatr Scand* **85**: 440–443.

Lebowitz BD, Pearson JL, Schneider LS. 1997. Diagnosis and treatment of depression in late life: consensus statement update. *JAMA* **14**: 1186–1190.

Lee HCB, Chiu HFK, Kwok WY. 1993. Chinese elderly and the GDS short form: a preliminary study. *Clin Gerotologist* **14**: 37–42.

Lyness JM, Noel TK, Cox C. 1997. 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* **157**: 449–454.

Mui AC. 1996. Geriatric depression scale as a community screening instrument for elderly Chinese immigrants. *Int Psychogeriatr* **8**: 445–458.

Pomeroy IM, Clark CR, Philp I. 2001. Short scales were as effective as long scales in screening for depression in older patients. *Int J Geriatr Psychiatry* **16**: 321–326.

Reynolds CF. 1996. Depression: making the diagnosis and using SSRIs in the older patient. *Geriatrics* **51**: 28–34.

Rihmer Z, Angst J. 2005. Mood disorders: epidemiology. In: *Comprehensive Textbook of Psychiatry. 8th edn.* Sadock BJ, Sadock VA (eds). Lippincott: New York. Williams & Wilkins; p. 1575–1581.

Robison J, Gruman C, Gaztambide S. 2002. Screening for depression in middle-aged and older Puerto Rican primary care patients. *J Gerontol* **57A**: 308–315.

Schreiner AS, Hayakawa H, Morimoto T. 2003. Screening for late life depression: cutoff scores for the Geriatric Depression Scale and the Cornell Scale for Depression in Dementia among Japanese subjects. *Int J Geriatr Psychiatr* **18**: 498–505.

Shahmohammadi D. 1990. *Integration of Mental Health in Primary Health Care (PHC) in Rural Areas of Shahrekord. Project Final Report*. Tehran Psychiatric Institute.

Sheikh JI, Yesavage JA. 1986. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol* **5**: 165–173.

Steffens DC, Skoog I, Norton MC. 2000. Prevalence of depression and its treament in an elderly population: the Cache County study. *Arch Gen Psychiatry* **57**: 601–607.

The EE. Validation of Malay version of Geriatric Depression Scale among elderly inpatient. Available from: URL: http://www.priory.com/psych/MalayGDS.htm

WHO, Department of Chronic Disease and Health Promotion (CHP): Aging and life course. Available from: URL: http://www.Who.int/hpr/aging/index.html

World Health Organization. 1990. *Composite International Diagnostic Interview, version 1*. WHO: Geneva.

Yesavage JA, Brink TL. 1983. Development and validation of a Geriatric Depression Scale: a preliminary report. *J Psychiar Res* **17**: 37–49.