Problems with the Japanese Version of the Center for Epidemiologic Studies Depression (CES-D) Scale when Applied to Japanese Independent Community-Dwelling Elderly: an Examination of the Factor Structure

Shinichi DEMURA¹ and Susumu SATO²

¹Faculty of Education, Kanazawa University, Ishikawa, Japan

²Life-long Sports Core, Kanazawa Institute of Technology, Ishikawa, Japan

Abstract

Purpose: The present study aimed to determine the problems of the Japanese version of the Center for Epidemiologic Studies Depression (CES-D) scale when applied to Japanese independent community-dwelling elderly focusing on its factor structure.

Methods: The subjects were 1791 community-dwelling independent elderly aged 60 and over (848 males and 943 females). This study used the Japanese version of the CES-D scale to assess depression in the elderly.

Results: From the results of an exploratory factor analysis, the four factors of somatic symptoms, depression affect, positive affect and relation with others were interpreted. These factors were the same as those in the previous study, but the items making up each subscale differed. Therefore, the construct validity of each subscale was not necessarily assured. In the confirmatory factor analysis, goodness-of-fit was high for both the first and second-order factor models. In examining the validity by Akaike's information criterion (AIC), the second-order factor model assuming depression as a higher-order factor among the four factors was a better fit than the first-order factor model. Although the reliability of each subscale was not sufficient, adequate reliability was assured in the total scale.

Conclusion: An assessment of depression using the Japanese version of the CES-D scale should be conducted using the total score, while a re-examination of items making up each subscale is needed.

Key words: depression, center for epidemiologic studies depression (CES-D) scale, geriatric depression scale (GDS), exploratory factor analysis, confirmatory factor analysis

Introduction

With an advancing and aging society, the importance of the assessment of mental health in the elderly is increasing. Depression is one of the more important psychological characteristics of mental health, and the assessment of depression in the elderly needs to be carried out properly.

Previous studies have developed several depression scales, such as the self-rating depression scale (1), the geriatric depression scale (GDS) (2), Hamilton's depression scale (3), and the Center for Epidemiologic Studies Depression (CES-D) scale (4). The validity of the CES-D scale as a psychological scale

was verified in clinical field trials. This scale is one of the major depression scales used around the world (5). The CES-D scale was developed by the National Institute of Mental Health (NIMH) to screen for symptoms of depression in the ordinary population (6). In Japan, Shima et al. (7) translated the CES-D scale into Japanese, and this scale has been used in many applied studies (8–12).

The CES-D scale consists of 20 items, and assesses depression symptoms from the total score. In clinical settings, screening for depression symptoms is generally carried out using a cut-off value of 16 points of the total score (4, 6). Recently, by using the subscales of the CES-D scale, a more detailed evaluation of depression in the elderly has been attempted (7). The modeling of subscales in the CES-D scale has already been examined in developing this scale, and the four subscales, depression affect, positive affect, somatic symptoms and decline of activity, and relationship with others, were interpreted (4). Recently, however, three factor structures for Chinese-Americans and two factor structures for Chinese living

Received Apr. 1 2003/Accepted Sep. 12 2003

Reprint requests to: Susumu SATO

Life-long Sports Core, Kanazawa Institute of Technology, 7-1, Ohgigaoka, Nonoichi, Ishikawa 921-8501, Japan

TEL: 076-248-1100, FAX: 076-294-6701

E-mail: sssato@neptune.kanazawa-it.ac.jp

in Hong Kong have been reported (13). Similar results have been reported in previous studies (14–16). Using the Japanese version of the CES-D scale, although Yatomi et al. (12) interpreted four factors which were equal to Radloff's result (4), they reported that goodness-of-fit in their confirmatory factor analysis was not necessarily high, and that further research was needed. Iwata et al. (8–11) reported that differences in factor structure and internal consistency for the CES-D scale were found among different age groups of Japanese samples and among different ethnocultural populations. These reports suggest the possibility that there are differences in the factor structure of depression based on culture and race, and that similar assessment procedures were not always appropriate when translating the original version into other languages and applying it to different ethnocultural populations.

In Japan, the existing depression scales developed in other countries have often been used to assess depression symptoms of the elderly. Therefore, it is important to confirm the reliability, factor structure and validity of the Japanese version of the CES-D scale, and to examine whether assessment by total or subscale score is appropriate. Furthermore, many of the previous studies concerning this scale, such as that by Iwata et al. (8–11), focused on young or moderate populations, and there are few studies focusing on the elderly population and determining the problems applying to them.

The present study aimed to examine the factor structure of the CES-D scale when applied to independent community-dwelling elderly Japanese living in the local area, and to verify its reliability and validity as a depression scale. These examinations determined some problems with the Japanese version of the CES-D scale and provide indications as to whether or not it is appropriate for assessing depression symptoms in elderly Japanese, and whether or not development of another depression scale for elderly Japanese is needed.

Methods

Subjects and data collection

We conducted a survey at Kushiro City in Hokkaido, Akita City in Akita, Mito City in Ibaragi, Kanazawa City in Ishikawa, Fukui City in Fukui, Komaki City in Aichi and Gifu City in Gifu based on the limitations of availability within our research fields. We delivered a questionnaire to a total of 1900 subjects (100 to 300 questionnaires were distributed to each area), and collected data from 1865 subjects. After confirming the data loss for gender, age and disease history, data from 1791 subjects (848 males, 60s: n=134, 65s: n=233, 70s: n=249, 75s: n=139, 80s: n=93; and 943 females, 60s: n=166, 65s: n=247, 70s: n=238, 75s: n=167, 80s: n=125) were used for the analysis (effective response rate 94%). The investigators in this study were researchers working at universities in each prefecture. They gathered data in a general delivery survey and interview setting or education class settings. Prior to the survey, investigators explained to the subjects that they could refuse to participate in the survey, and that participating would not adversely affect their privacy.

Subjects were community-dwelling independent elderly people aged 60 and over, almost all of whom participated in

health or culture education classes for the elderly offered by each municipality, such as healthy exercise programs, recreational sports and traditional culture programs on flower arrangement, calligraphy, "go" games and haiku poems.

Questionnaire

This study used the CES-D scale to assess depression in the elderly. The scale was translated into Japanese by Shima et al. (7), and consists of 20 items with a 4-point Likert scale from never=0 to always=3 (for four items, the pattern of the reverse rating scale was from never=3 to always=0). The depression level advanced with increasing item score.

In addition, we also used the GDS. This scale has been used around the world due to its simplicity. The GDS consists of 15 items with a dichotomous scale.

Statistical analyses

Confirmation of the factor structure of the Japanese version of the CES-D scale by exploratory factor analysis

The four-factor structure of the original version and the Japanese version of CES-D scale have been reported by Radloff et al. (4), Shima et al. (7) and Yatomi et al. (12). However, previous studies reported that an equal factor structure was not always obtained when applied to different ethnocultural populations. Therefore, this study attempted to confirm whether the factor structure comparable to previous studies is interpreted by exploratory factor analysis when applied to Japanese elderly. Exploratory factor analysis was applied to the correlation matrix consisting of 20 items of the CES-D scale. Extraction of factors was conducted by considering the following: magnitude of eigenvalue (1.0 or more), scree-plot, and interpretability of factors.

Examining the goodness-of-fit of factor model of the Japanese version of CES-D by confirmatory factor analysis

In assessment of depression symptoms with the CES-D scale, two assessment procedures using total score and subscale score have been examined. This means that two different models of the first-order factor model, which comprises the four factors interpreted in previous studies (4, 7), and the second-order factor model composed of a higher-order factor (depression) and four sub-factors, are assumed. When applying the Japanese version of the CES-D scale to Japanese elderly, it should be determined which assessment procedure (factor model) is more appropriate. Therefore, this study examined the goodness-of-fit of these two factor models for Japanese elderly by confirmatory factor analysis. The model-fitting was evaluated by calculating the goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI) and Akaike's information criterion (AIC).

Reliability of total and sub-scales

For each factor interpreted in this study, the reliability of each subscale items was examined using the Cronbach alpha coefficient. In addition, the relationships between the subscale scores of the Japanese version of the CES-D scale were calculated based on the factor structure of Yatomi et al. (12), and those based on the factor structure extracted in this study were examined by Pearson's correlation coefficients.

Relationships with other depression scale scores

The relationships between the score on the GDS and the CES-D subscale scores extracted in this study were examined by Pearson's correlation coefficients.

Results*Confirmation of factor structure of the Japanese version of the CES-D scale by exploratory factor analysis*

Table 1 shows descriptive statistics for the CES-D scale scores. The mean value of the total score of the CES-D scale was 9.9, and was equal to that of the normal adult group in the

Table 1 Descriptive statistics of CES-D scale

	N	Mean	SD	Skewness	Kurtosis
1 I was bothered by things that usually don't bother me	1195	0.59	0.56	0.28	-0.84
2 I did not feel like eating; my appetite was poor	1198	0.28	0.50	1.59	1.64
3 I felt that I could not shake off the blues even with help from my family and friends	1186	0.26	0.50	1.77	2.30
4 I felt that I was just as good as other people	1161	0.79	0.61	0.14	-0.49
5 I had trouble keeping my mind on what I was doing	1174	0.47	0.60	0.90	-0.19
6 I felt depressed	1165	0.43	0.58	1.00	0.01
7 I felt like everything I did was an effort	1177	0.63	0.62	0.46	-0.66
8 I felt hopeful about the future	1122	0.95	0.49	-0.12	1.14
9 I thought my life had been a failure	1153	0.33	0.54	1.41	1.03
10 I felt fearful	1169	0.32	0.55	1.46	1.19
11 My sleep was restless	1182	0.58	0.64	0.65	-0.56
12 I was happy	1163	0.80	0.50	-0.35	0.14
13 I talked less than usual	1169	0.45	0.63	1.07	0.06
14 I felt lonely	1269	0.70	0.46	-0.88	-1.23
15 People were unfriendly	1158	0.29	0.53	1.65	1.80
16 I enjoyed life	1167	0.74	0.55	-0.06	-0.41
17 I had crying spells	1169	0.31	0.52	1.44	1.12
18 I felt sad	1167	0.37	0.57	1.22	0.50
19 I felt that people disliked me	1179	0.22	0.46	1.90	2.80
20 I could not 'get going'	1187	0.55	0.62	0.68	-0.51
Total score	1105	9.91	5.11	0.74	0.47

Table 2 The result of exploratory factor analysis for 20 items of the Japanese version of CES-D scale

No.	Contents	F1	F2	F3	F4	Communality
18	I felt sad	0.821	-0.004	-0.026	-0.108	0.565
17	I had crying spells	0.773	-0.143	-0.058	0.003	0.473
3	I felt that I could not shake off the blues even with help from my family and friends	0.435	0.142	0.038	0.129	0.408
14	I felt lonely	-0.429	-0.035	-0.030	-0.082	0.265
1	I was bothered by things that usually don't bother me	0.396	0.244	-0.009	-0.022	0.321
6	I felt depressed	0.377	0.311	0.004	0.078	0.462
10	I felt fearful	0.334	0.167	-0.007	0.096	0.278
7	I felt like everything I did was an effort	-0.070	0.813	0.035	-0.059	0.555
5	I had trouble keeping my mind on what I was doing	-0.006	0.664	-0.011	0.030	0.458
20	I could not 'get going'	-0.005	0.660	0.029	0.004	0.443
2	I did not feel like eating; my appetite was poor	0.141	0.410	-0.041	-0.111	0.195
13	I talked less than usual	-0.089	0.355	0.031	0.330	0.314
11	My sleep was restless	0.294	0.319	-0.018	-0.038	0.278
12	I was happy	0.044	-0.081	0.761	0.008	0.572
16	I enjoyed life	0.051	-0.022	0.684	-0.001	0.475
8	I felt hopeful about the future	-0.038	0.031	0.483	-0.051	0.229
4	I felt that I was just as good as other people	-0.145	0.125	0.337	-0.015	0.127
15	People were unfriendly	-0.027	-0.085	-0.032	0.767	0.488
19	I felt that people disliked me	0.084	0.007	-0.049	0.639	0.479
9	I thought my life had been a failure	0.236	-0.003	0.066	0.309	0.257
F1	Depression affect	1.000				
F2	Somatic symptoms	0.638	1.000			
F3	Positive affect	0.153	0.188	1.000		
F4	Relation with others	0.633	0.601	0.183	1.000	

Exploratory factor analysis based on maximum-likelihood method and promax rotation was conducted.

previous study. The proportion of subjects showing depression symptoms (scoring 16 or more points used as the cut-off point for depression screening) was 14.3%.

Table 2 shows the factor pattern matrix when the exploratory factor analysis is applied to the correlation matrix from the 20 subscale items of the CES-D scale. The four factors of depression affect, positive affect, somatic symptoms and relations with others, which were the same as that of the previous study, were interpreted.

Examining the goodness-of-fit of factor model of the Japanese version of the CES-D scale by confirmatory factor analysis

Figures 1 and 2 show model-fitting when assuming the first-order factor model and the second-order factor model for the 20 items of the CES-D scale, respectively. In both models, the GFI and AGFI values were high, and good model-fitting was found. The goodness-of-fit indices and AIC fitted the

second-order factor model better than the first-order factor model.

Reliability of total and subscale of the Japanese version of the CES-D scale

Table 3 shows Cronbach's alpha coefficients for the total scale and subscales. Except for physical symptoms (0.747), the coefficients of the subscales were all lower than 0.647. The alpha coefficient of the total scale was 0.813. Furthermore, the correlations between factor scores based on two factor structures obtained in Yatomi et al.'s reports (12) and in this study were calculated. These correlations were significant and high (0.896 to 1.000, $p < 0.01$).

Relationships with other depression scores

As shown in Table 4, the correlation coefficients between the CES-D scale subscale scores obtained in this study and the total score of the GDS were significant (0.239 to 0.579, $p < 0.01$).

Fig. 1 Model fitting for the first-order factor model
GFI: 0.933, AGFI: 0.914, AIC: 1089.94

Fig. 2 Model fitting for the second-order factor model
GFI: 0.953, AGFI: 0.941, AIC: 712.78

Table 3 The result of examining reliability and reproducibility of factors

Subscales	No.	Item contents	alpha	r#
Depression affect	1	I was bothered by things that usually don't bother me	0.644	0.948
	3	I felt that I could not shake off the blues even with help from my family and friends		
	6	I felt depressed		
	10	I felt fearful		
	14	I felt lonely		
	17	I had crying spells		
	18	I felt sad		
Somatic symptoms	2	I did not feel like eating; my appetite was poor	0.747	0.957
	5	I had trouble keeping my mind on what I was doing		
	7	I felt like everything I did was an effort		
	11	My sleep was restless		
	13	I talked less than usual		
Positive affect	20	I could not 'get going'	0.647	1.000
	4	I felt that I was just as good as other people		
	8	I felt hopeful about the future		
	12	I was happy		
Relation with others	16	I enjoyed life	0.616	0.896
	9	I thought my life had been a failure		
	15	People were unfriendly		
	19	I felt that people disliked me		
Total scale			0.813	

alpha: Cronbach's alpha coefficient, r#: Correlations between factor scores calculating based on factor structures obtained in Yatomi et al.'s report (12) and in this study.

Table 4 The relationship between CES-D total and subscale scores and GDS score

	r
Depression affect	0.519
Somatic symptoms	0.579
Positive affect	0.239
Relation with others	0.417
Total scale	0.344

r indicates the correlations between total and subscale scores of CES-D scale obtained in this study and total score of GDS.

Discussion

The sample of this study was independent Japanese elderly. Seven percent of them lived alone and the others lived with their spouse, child or grandchild. In the self-evaluation of health status, 62% responded "average," and 12% responded "bad" or "very bad." In participation in social or leisure activities, 80% responded "do" or "do sometimes." In depression symptoms, the proportion of subjects showing depression symptoms (scoring 16 or more points used as the cut-off point for depression screening) was 14.3%. In Shima et al.'s report (6) examining the validity of the Japanese version of the CES-D scale, the proportion of subjects in a normal control group showing 16 or more points was 15.2%, and the value for the group aged 60 or more reported by Clark et al. (17) was 16.2%. These values are considered to be almost comparable to this study. Under the limitation of these characteristics of sample, this study examined the problems of the Japanese version of the CES-D scale when applied to elderly Japanese from the viewpoint of the characteristics of the factor structure of this scale.

In the CES-D scale frequently used in clinical settings, total score is used as an assessment scale to screen for depression symptoms (6). Although the factor structures of the original and Japanese version of the CES-D scale were reported in the previous studies of Radloff (4) and Yatomi et al. (12), such as the above four-factor structure, this is merely a classification of depression symptoms. Furthermore, the utility of each subscale as a depression assessment scale has not been sufficiently examined.

Therefore, this study conducted a confirmatory factor analysis of 20 items of the Japanese version of the CES-D scale to determine the reproducibility of the factor structure. As a result, the four factors of depression affect, somatic symptoms, positive affect and relation with others were extracted and were the same as those in the previous study. The items included in each factor were the same as those of Yatomi et al. (12) except for two items of No. 1, "bothered by things," and No. 9, "life had been a failure." "Bothered by things" belonged to the somatic symptoms factor in Yatomi et al.'s report (12), but to the depression affect factor in this study. "Life had been a failure" belonged to the depression affect factor in Yatomi et al. (12), but to the relation with others factor in this study. The relationships between subscales based on Yatomi et al.'s factor structure and those based on the factor structure of this study were high (Table 3). It is thought that the reproducibility and generalization of this four-factor structure is high as the factor structure of the Japanese version of the CES-D.

On the other hand, because the relationships between factors were also high, it is thought that the factor structure of the Japanese version of the CES-D scale is not clear and simple. Therefore, although it is considered valid to interpret these four factors as subscales of the Japanese version of the CES-D scale, a more detailed assessment of depression by simply scoring

subscales may not be appropriate. When developing a psychological scale using factor analysis, the factor structure of the scale should be examined after determining the validity of the subscale items using a stepwise variable selection in an exploratory factor analysis (18). The Japanese version of the CES-D scale, however, was developed under the restriction that the number and content of the subscale items not be changed from the original version (7, 12). This may also be one of the reasons that a few items showing low internal consistency were included among the subscale items. Therefore, reselection of the depression items, including a change of items, may be required to develop a depression scale with high factor validity.

In general, the establishment of a model is useful to our understanding of psychological phenomena. In structural equation modeling in multivariate analysis, there are factor models such as a first-order factor model and a hierarchical factor model. Recently, a higher-order factor model and a multiple indicator multiple case (MIMIC) model have been proposed in covariance structure analyses (19). In the previous study determining the modeling of depression in the elderly, a second-order factor model assuming a higher-order factor and a first-order factor model have been examined (7, 12). We examined the validity of these models by confirmatory factor analyses. Goodness-of-fit indices were high in both models. Comparing the AIC values of the two models, the second-order factor model was considered to be more reliable than the first-order factor model. These results indicate that although generalization of the four factors is almost valid in the Japanese version of the CES-D scale, more appropriate model-fitting can be obtained by assuming a higher-order factor for depression as a general factor.

In general, the level of reliability coefficient (Cronbach's alpha coefficient) of a psychological questionnaire is required to be 0.8 or more (20). The reliability coefficients of the four subscales in the CES-D scale were lower than 0.8, and hence their levels were not sufficiently high. Especially, the number of items making up the subscale of relation with others was minimal (3 items), and insufficient to perform a reliable and stable assessment. Therefore, it may be problematic to use the subscale of relation with others as an independent assessment scale of depression in the elderly. In addition, considering the

result of examining factor validity by confirmatory factor analysis, re-examination of the subscale items for the relation with others factor is also a problem that we should solve. On the other hand, the reliability coefficient of the total scale was 0.813, and the reliability of the Japanese version of the CES-D scale is considered to be assured in a general assessment by total score.

In the relationship between the Japanese version of the CES-D scale and the GDS, which is a general-purpose depression scale for the elderly, the correlations between the CES-D subscale scores and the GDS total score varied (0.239 to 0.570), and it was considered that these two depression scales do not always assess a common content of depression. The GDS assumes a single factor structure, and the total score is used for assessment. The predominance of the CES-D scale over other depression scales is due to the fact that this scale can assess depression with a multidimensional scale. Therefore, the factor validity of the subscale items and the reliability of each subscale are important problems that should be addressed.

In conclusion, in the exploratory factor analysis, the four factors of depression affect, somatic symptoms, positive affect and relation with others were interpreted. These factors were the same as those in the previous study, but each subscale item partly differed, so that the factor validity of each subscale was not necessarily assured. In addition, although the reliability of each subscale was not sufficient, sufficient reliability was assured in the total scale. In the confirmatory factor analyses for the first-order and second-order factor models, goodness-of-fit was high in both models. In examining the validity by AIC, the second-order factor model assuming a higher-order factor for depression fitted better compared with the first-order factor model. As mentioned above, assessment of depression by the Japanese version of the CES-D should be conducted using the total score, and re-examination of the individual items of each subscale is needed.

Acknowledgments

This research was supported in part by a Grant-in-Aid for Scientific Research, the Japan Ministry of Education, Science, Sports and Culture (2000–2002 Demura Project #12680019).

References

- (1) Zung WW, Richards CB, Short MJ. Self-rating depression scale in an outpatient clinic. Further validation of the SDS. *Arch. Gen. Psychiatry* 1965; 13: 508–515.
- (2) Gerety C. Medical evaluation of the geriatric patient. In: *Geriatric Medicine* ed. by Cats MS, Churchill Livingstone Publication, Appendix D, 1982; p. 31.
- (3) Hamilton M. A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* 1960; 23: 56–62.
- (4) Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl. Psychol. Measurement* 1977; 1: 385–401.
- (5) Sugisawa H, Shibata H. Psychosocial determinants of changes in activities of daily living and depressive status among stroke patients at home. *Nippon Koshu Eisei Zasshi* 1995; 42: 203–209. Japanese.
- (6) Shima S. CES-D Scale. Chiba test center 1998; Tokyo, pp. 1–12. (in Japanese).
- (7) Shima S, Shikano T, Kitamura T. A new self-report depression scale. *Psychiatry* 1985; 27: 717–723. (in Japanese).
- (8) Iwata N, Roberts RC, Kawakami N. Japan-U.S. comparison of responses to depression scale items among adult workers. *Psychiatry Res.* 1995; 58: 237–245.
- (9) Iwata N, Roberts RE. Age differences among Japanese on the Center for Epidemiologic Studies Depression Scale: an ethnocultural perspective on somatization. *Soc. Sci. Med.* 1996; 43(6): 967–974.
- (10) Iwata N, Umesue M, Egashira K, Hiro H, Mizoue T, Mishima N, Nagata S. Can positive affect items be used to assess

- depressive disorders in the Japanese population? *Psychol. Med.* 1998; 28(1): 153–158.
- (11) Iwata N, Buka S. Race/ethnicity and depressive symptoms: a cross-cultural/ethnic comparison among university students in East Asia, North and South America. *Soc. Sci. Med.* 2002; 55(12): 2243–2252.
 - (12) Yatomi N, Liang J, Krause N, Akiyama H. Measurement of depression symptoms of the Japanese elderly by the CES-D scale: examination of factor structure and cultural differences. *J. Social Gerontology* 1993; 37: 37–47.
 - (13) Cheung CK, Bagley C. Validating an American scale in Hong Kong: the Center for Epidemiological Studies Depression Scale (CES-D). *J. Psychol.* 1998; 132: 169–186.
 - (14) Kuo WH. Prevalence of depression among Asian-Americans. *The Journal of Nervous and Mental Disease* 1984; 172: 449–457.
 - (15) Liang, J, Tran TV, Krause N, Markides KS. Generational differences in the structure of the CES-D scale in Mexican Americans. *Journal of Gerontology* 1989; 44: S11–120.
 - (16) Guarnaccia PJ, Angel R, Worobey JL. The factor structure of the CES-D in the Hispanic health and nutrition examination survey: the influences of ethnicity, gender and language. *Soc. Sci. Med.* 1989; 29: 85–94.
 - (17) Clark VA, Aneshensel CS, Freichs RR, Morgan TM. Analysis of effects of sex and age in response to items on the CES-D scale. *Psychiatry Res.* 1981; 5: 171–181.
 - (18) Kano Y, Harada A. Stepwise variable selection in factor analysis. *Psychometrika* 2000; 65: 7–22.
 - (19) Toyoda H. Introductory Analysis of covariance structures. Asakura Shoten 1998; Tokyo, pp. 56–66. (in Japanese).
 - (20) Shigemasu K. Psychometric testing: The society for the Promotion of the university of the Air 1999; Tokyo, pp. 76–79.