

A Validation Study on the Filipino Geriatric Depression Scale (GDS) using Rasch Analysis

Welison Evenston G Ty^{1*}, Roger D Davis¹, Maria Isabel E Melgar¹ and Miguel A. Ramos Jr.²

¹Ateneo de Manila University, Quezon City, Philippines.

²St. Luke's Medical Center, Philippines.

*Correspondence:

Welison Evenston G. Ty, Department of Psychology, Ateneo de Manila University, Quezon City, Philippines.

Received: 26 October 2019; Accepted: 18 November 2019

Citation: Welison Evenston G Ty, Roger D Davis, Maria Isabel E Melgar, et al. A Validation Study on the Filipino Geriatric Depression Scale (GDS) using Rasch Analysis. Int J Psychiatr Res. 2019; 2(7): 1-6.

ABSTRACT

At present, there is a dearth of studies that assess the “cultural equivalence” and dimensionality of the Filipino translated Geriatric Depression Scale (GDS) using item response theory methods. Using the Rasch model, this study sought to evaluate a Filipino-translation of GDS items using two parameters--one for ability to each person, and another for difficulty to each item. Doing so overcomes a source of continuing controversy associated with Likert scales. A total of 505 elderly respondents participated in the study. Results provided support that the 30 Filipino translated items form a unidimensional measure of depressive mood, as shown by their acceptable fit statistics; and hence, can be used in research as a valid, interval-level measure of depression. Subsequent Rasch analyses performed also revealed that 18 of 30 items were functioning equally across the two languages, and, of the 18 items, 3 items that assessed anxiety were dropped thereby creating a Filipino GDS Short Form (GDS-SF-F) that obeyed the requirements of the Rasch model, while conforming to the content of depression as given in the DSM-5. The study concludes that the 15-item GDS-SF-F can be used as a short measure of depression that provides a genuine interval level of measurement, with adequate internal consistency.

Keywords

Geriatric Depression Scale (GDS), Geriatric Depression Scale Short Form (GDS-SF), Rasch Analysis, Scale Translation, Scale Validation.

The Philippines has a population of over 100 million, an estimated 7% of whom are elderly people over 60 years [1]. This slice of the population represents 7 million men and women who will need medical and psychosocial care as they add years to their lives. Advances in diagnostic testing and imaging have made it possible to assess physiological conditions in the elderly with sophisticated accuracy and clarity. Likewise, the development of manuals such as the Diagnostic and Statistical Manual of Mental Disorders (DSM) and International Classification of Diseases (ICD) have operationalized the content of mental disorders, leading to increases in the reliability of diagnosis. Nevertheless, the diagnostic criteria themselves sometimes reflect not only changes in mental states, but also changes that accompany aging.

The diagnosis and assessment of depression and negative mood states provide an example. Depression is widely seen as the most

common psychiatric disorder among the elderly [2-3]. Additional research has shown that depression is often under-recognized and under-treated [4-6]. The medical condition of aging patients and their multiple medications may combine to mask the existence and severity of depression [7]. In one study of patients with multiple illnesses, the geriatrician was not able to identify depression in more than half the sample [8]. While, in other study, elderly patients were found to consult physicians mostly for physiological symptoms rather than psychological concerns [9].

A popular screening instrument used for depression in the elderly is the Geriatric Depression Scale (GDS) [10]. The GDS was developed in order to focus on depressive mood and to minimize the physiological effects of depression that form the second factor of the Beck Depression Inventory [11]. As a result, the instrument is more valid with an older sample where physiological problems become a confounding factor in the assessment of depression. Depression is often manifested through reduced energy and concentration, sleep problems, decreased appetite, weight loss and body complaints [2] which may be manifested by the elderly as either the normal effects of aging or the results of physical illness [10].

The long form version of the GDS is a dichotomous self-rating scale with thirty items intended primarily for use in clinical settings. In a review of the psychometric properties of the GDS, Yesavage et al. [10] reported an alpha coefficient of .94 in a mixed sample of unaffected and elderly people with depression. Kiefer and Reese [12] further reviewed 338 studies that used the GDS, 35 of which reported more than one type of reliability coefficient, and a mean internal consistency of .83. Since the Cronbach's alpha and the reliability coefficient are strongly influenced by scale length, this value is not unexpected for a scale consisting of 30 items [11].

The 30-item GDS has been translated into several languages, including Malay [13], Nepalese [14], Korean [15], Iranian [16] and Chinese [17]. The GDS was also translated to Filipino [18], with a Cronbach's alpha of .87 and a test-retest reliability of .74. A short form of the GDS was created by Sheikh and Yesavage [19] to reduce fatigue [20], especially among patients suffering from physical illness or patients with dementia. The GDS Short Form (GDS-SF) retained the 15 items with the highest item-total correlations from the original 100-item pool of the GDS. The GDS-SF consists of items which revolve around life satisfaction and happiness, enthusiasm about new activities, interest, energy, hopelessness and helplessness and fear about the future. Since the abbreviated form measures subjective aspects of depression, its main function is for depression screening rather than for diagnostic classification [21]. Scores between 0 and 4 are considered normal whereas scores ranging from 10 to 15 indicate moderate to severe depression. Several authors have reported correlations between scores on the short and long forms: .84 [19], .89 [22] and .66 in patients without severe depression [23].

The standard 15-item GDS-SF has also been translated into various languages (e.g., 16, 24, 25). Both categorical and item-response theory (IRT) methods have been used to evaluate the validity of these translations. Categorical methods have the virtue of establishing the sensitivity and specificity of the GDS-SF against established diagnostic systems, resulting in sensitivity and specificity values for their respective sites [16,26-28].

In contrast, IRT methods allow the dimensionality of scale items to be assessed (e.g., 29, 30). Items may be evaluated in terms of their fit statistics. Poorer fitting items may need to be rewritten, or may represent attributes not strongly associated with the construct. Chiang, Green, and Cox [21] used Rasch analysis to combine categorical and dimensional approaches to validity. A total of 177 U.S. older adults participated in the study that examined scale dimensionality, reliability, invariance, continuity, cutoff scores and diagnostic use of GDS-SF. About one third of the items were found to have different meanings for different groups of participants. Nevertheless, the GDS-SF was found to discriminate between older adults who have a moderate level of depression than between those with mild depression and no diagnosed depression.

The present study continues the effort to evaluate Filipino translations of the GDS items, in this case using the Rasch model. The Rasch model is the simplest of the latent trait models. In

classical psychometrics, the items have no difficulty parameter. The Rasch model adds a parameter for ability to each person, and a parameter for difficulty to each item. This overcomes a common objection and source of continuing controversy associated with Likert scales, namely that the summation of scores on Likert items constitutes ordinal, rather than interval level data.

From a psychiatric perspective, items which are easier (i.e., lower difficulty level) are likely to represent symptoms that are the first to emerge in the expression of a syndrome. Items which are harder (i.e., higher difficulty level) represent symptoms that are likely to emerge as the syndrome becomes more severe. For example, guilt can reach psychotic proportions in depressed patients, but this symptom is expected to emerge as an expression of severe depression, not mild depression.

Because the items occur in a particular order, they are said to form an "item hierarchy." Construct validity in the context of the Rasch model thus has a different meaning than it does in classical psychometrics. In classical psychometrics, all items are "on the same level." As such, the foundation of construct validity is said to be content validity. That is, "does the scale measure everything that should be measured?" In the Rasch model, the item difficulty parameter requires that items be inspected with regard to their order. That is, "does the manifestation of the construct unfold in a manner that accords with the theory of the construct?"

Translation of instruments across cultures and languages can be problematic for a number of reasons. First, a construct may not be equivalent across cultures. The construct may exist across cultures, but have features unique to one culture that do not exist in other cultures. Second, due to cultural stigma, some groups may be very reluctant to report symptoms of psychopathology, but find it perfectly acceptable to mention fatigue or boredom. Such translated items may form a Rasch scale, but be changed in terms of their difficulty level, thereby influencing the item hierarchy. For example, an item such as "I feel down" could be translated to mean "I feel suicidal." Another item such as "I feel hopeless" could be translated to mean "I feel pessimistic." The severity of the item has obviously changed, affecting the item hierarchy. Where the difficulty level of the translated items is found to vary across cultures, there is always the question concerning whether this difference is caused by the translation of the items or by differences in the construct. Only judgment can answer this question. To date, no local studies have assessed the "cultural equivalence" and dimensionality of the Filipino translated GDS using item response theory methods.

A single existing study [18] factor analyzed the Filipino GDS items, resulting in one dominant factor and one minor factor. The said research also found that the GDS possessed strong internal consistency yielding a Cronbach's alpha of .87; a test-retest reliability within a span of 2 weeks of $r = .74$ ($p < .01$); and an odd-even reliability and split-half reliability coefficients of $r = .71$ ($p < .01$) and $r = .74$ ($p < .01$) respectively.

Method

Participants

A total of 505 elderly respondents gave informed consent to participate in the study. Participants were required to be age 60 or older, to be able to comprehend both written and verbal English and Filipino, and to exhibit no evidence of cognitive impairment. Respondents were required to complete all items from the English and the Filipino versions of the GDS. A total of 45 respondents who failed to complete all items were dropped. This included 24 participants with an incomplete English version and 21 participants with an incomplete Filipino version. The majority of the elderly respondents were outpatients ($n = 416$). A minority were recruited from the geriatric care centers in tertiary hospitals ($n = 89$).

Of the 505 elderly participants, 250 answered the English version of the GDS [10] while 255 answered the Filipino translation of the GDS. Of the 250 respondents of the English version of the GDS, 12 were acquired through the geriatric care centers of tertiary hospitals while 238 were out-patients. Of the 255 respondents of the Filipino version of the GDS, 65 were acquired from the geriatric care centers in tertiary hospitals while 190 were out-patients. Any respondent for whom there was a question of possible depression was included, thereby ensuring that both subthreshold and severe forms of the disorder would be represented in the analysis. We did not attempt to establish a formal diagnosis, simply because we could not be sure both the items and the diagnostic procedures would translate across languages. For this reason, we did not seek to equate the groups in terms of their level of depression. Instead, we sought enough data to study the item hierarchies that emerged in the English and Filipino items.

Procedure

During the initial phase of the study, a research team composed of 1 geriatric specialist, 2 psychologists, and a psychiatrist was tasked to translate the items of the long form of the GDS. We chose the long form of the instrument because we were aware that a certain proportion of the items might not translate successfully. We wanted to ensure that if translated items needed to be dropped, other items assessing similar content would be available, thereby preserving the content validity of the scale.

Item translation was conducted in relation to the research team's knowledge of local elderly population and academic background in psychological assessment. The exact wording of the translations was worked out by panel discussions of the research team. The provisional items were then given to a Filipino language consultant for review and comments. After this, the research team deliberated, evaluated, and revised the items again. A back translation was then performed to strengthen item translation. All items passed the back-translation process.

The second phase of the study involved the concurrent administration of the English and Filipino versions of the GDS to elderly participants. Members of the research team as well as graduate students in psychology administered the instrument to

consenting elderly participants. Administration of both versions of the GDS was researcher-aided. To complement the structured interview form used to gather data from respondents, members of the research team and graduate students who were involved in the administration were each armed with an 8.5 inch by 11 inch flip chart containing items from the GDS in large font size to aid administration. Members of the research team as well as the graduate students in psychology were also required to undergo an orientation to familiarize themselves with the administration process as well as to become accustomed to the administration protocol, the structured interview form, and the use of the flip chart provided. The orientation also provided a venue to discuss and address anticipated concerns and foreseen considerations that researchers may have concerning initial administration and re-administration of the GDS. These measures were taken in order to ensure the integrity of the data gathered. Retest of the English and Filipino version of the GDS was completed within a span of 14 days from the time of initial administration.

The items were then analyzed using the dichotomous version of the Rasch model, and tested for differential item functioning using the Mantel-Haenszel statistic in order to establish whether each item was successfully translated. Next, the Filipino items were Rasch analyzed alone, in order to determine whether the Filipino items would form a unidimensional depression scale on their own. This was done because translation may affect the difficulty of an item without changing its content validity. For example, the items "I am depressed" and "I am very depressed" obviously differ in their difficulty level. Yet, both assess depression. Such items would be flagged as significantly different by the Mantel-Haenszel statistic, yet possess acceptable fit statistics according to the Rasch model. Finally, we looked at the characteristics of the Filipino GDS-15, a short form consisting of 15 successfully translated items, 5 of which overlap the GDS-15 [19].

Results and Discussion

Descriptive Statistics

Scores on the GDS ranged from 0 to 22, with a mean of 6.27 ($SD = 4.67$). There was no significant difference in the mean geriatric depression rating of men ($M = 5.97$; $N = 89$) and women ($M = 6.56$; $N = 161$). We did not calculate the mean difference in depression scores among subjects completing the Filipino or English form, because doing so rests on the assumption that the items difficulties are, in fact, equivalent, and determining such equivalence was the purpose of the study.

Reliability

The Cronbach's alpha value for the GDS was .84, which indicates adequate internal consistency. Though lower values are seen in the literature, an alpha of .70 or greater is recommended [31]. Test-retest reliability of the scale was .91 when administered within a span of 14 days for the 30-item version of the GDS.

Rasch Analysis

Items were analyzed with the dichotomous version of the Rasch model. By convention, the mean of the item difficulties is set to 0

in the Rasch model. Negative values are indicative of items that are relatively easier than the “average item,” while positive values are indicative of items that are relatively harder. For the English items, difficulties ranged from -2.68 to 2.42, a total range of 5.1 logits.

The easiest items were “Do you prefer to stay at home, rather than go out and do new things?” at -2.68 logits and “Is your mind as clear as it used to be?” at -1.79 logits. The hardest items were “Do you feel that your situation is hopeless?” at 2.42 logits and “Do you think it is wonderful to be alive right now?” at 2.03 logits. Interpreted as an item hierarchy, this suggests that social isolation and lack of mental clarity tend to emerge first, and hopelessness tends to emerge last in the expression of the construct.

Outfit statistics for the English items ranged from .51 for “Do you feel that your situation is hopeless?” to 1.47 for “Do you feel hopeful about the future?” The purpose of the Rasch model is to transform ordinal-level data to interval-level data using the logistic function. In the Rasch model, each item is modeled to have an item characteristic curve (ICC) of the same slope and shape. As such, the only difference between the items is their level of difficulty. Fit statistics evaluate the extent to which the data for a particular item approximates this slope. Values below .50 indicate that the item functions like a switch, being almost completely passed above its difficulty level and almost completely failed below its difficulty level. Values beyond 2.0 indicate that the item is too noisy and degrades measurement. All items fell within the desired .50 to 1.5 range.

The mean ability of the person sample in the English items was -1.99. Since the mean of the item difficulties in the Rasch model is by convention 0, this indicates that the items were substantially more difficult than the subjects were depressed, a targeting issue. Nevertheless, the standard errors associated with the difficulty estimates across all items ranged only from .14 to .40 logits, indicating that enough information was available to estimate the difficulty of even the hardest item with considerable precision, due to the sample size. Item-scale correlations ranged from .17 to .51. Lower values appeared to be attenuated by low item endorsements. Cronbach’s alpha was .83. We do not report similar statistics for the Filipino items, because these items are understood to be provisional, and a few aberrant items would mischaracterize the pool.

In the next phase of the analysis, the English language and Filipino items were Rasch analyzed together. Differential item functioning was examined using the Mantel- Haenszel statistic [32] with a slice width of .1 logits for a total of 22 slices. Twelve of 30 items were significantly different in terms of their difficulty across the two groups at the $p < .05$ level. Four of the 12 items were significantly easier for the subjects completing the English form, while 8 of the 12 items were significantly easier for subjects completing the Filipino form. This suggests, then, that 18 of 30 items were functioning equally across the two languages. Table 1 presents these items in English and Filipino.

English Language GDS item	Filipino GDS item
4. Do you often get bored?	Madalas ba kayong mainip?
5. Are you hopeful about the future?	Kayo ba ay may nakikitang pag-asa sa kinabukasan?
6. Are you bothered by thoughts you can't get out of your head?	May gumugulo ba sa inyong isipan na hindi ninyo maalis?
8. Are you afraid that something bad is going to happen to you?	Nag-aalala ba kayo na may masamang mangyayari sa inyo?
9. Do you feel happy most of the time?	Madalas ba kayong masaya?
10. Do you often feel helpless?	Pakiramdam n'yo ba wala na kayong magawa sa inyong kalagayan?
11. Do you often get restless and fidgety?	Madalas ba kayong hindi mapalagay at hindi mapakali?
13. Do you frequently worry about the future?	Madalas ba kayong nag-aalala tungkol sa inyong kinabukasan?
16. Do you often feel downhearted and blue?	Madalas ba kayong malungkot at nasisiraan ng loob?
17. Do you feel pretty worthless the way you are now?	Sa kasalukuyan, pakiramdam n'yo ba nawawalan kayo ng halaga sa inyong sarili?
18. Do you worry a lot about the past?	Madalas ba kayong nag-aalala tungkol sa inyong nakaraan?
20. Is it hard for you to get started on new projects?	Nahihirapan ba kayong magsimula ng mga bagong gawain?
23. Do you think that most people are better off than you are?	Sa tingin n'yo ba mas nakakaangang ang masnakararami kumpara sa inyo?
24. Do you frequently get upset over little things?	Parati ba kayong nagagalit kahit sa maliliitna bagay lang?
26. Do you have trouble concentrating?	Nahihirapan ba kayong mag-“concentrate?”
27. Do you enjoy getting up in the morning?	Ginaganahan ba kayong bumangon sa umaga?
28. Do you prefer to avoid social gatherings?	Mas gusto n'yo bang umiwas sa mga pagtitipon?
29. Is it easy for you to make decisions?	Madali ba sa inyo ang gumawa ng desisyon?

Table 1: Successfully translated English and Filipino GDS items.

DSM-5 Criteria for Major Depressive Episode	GDS-Long	GDS-15	Filipino GDS
1. Depressed most of the day, nearly every day.	1,3,5,7,9,10,15,16,22,24,25,27	1,3,7,9,10,15,22	5,9,10,16,24,27
2. Diminished interest or pleasure	2,4,12,20,28	2,4,12	4,20,28
3. Weight loss or gain			
4. Insomnia or Hypersomnia			
5. Psychomotor agitation or retardation	11		11
6. Fatigue	19,21	21	
7. Worthlessness or guilt	17,23,8	8,17,23	8,17,23
8. Concentration issues or indecisiveness	14,26,29,30	14	26,29
9. Recurrent thoughts of death			
Anxiety (not in diagnostic criteria)	6, 13, 18		6, 13, 18

Table 2: GDS Forms Compared to DSM-5 Criteria for Major Depressive Episode¹.

¹Item 6 asks about “thoughts you can’t get out of your head,” but

does specify suicidality; it was therefore considered obsessive and classified as anxiety-related.

Analysis of the Filipino GDS Items

Because 12 of 30 Filipino items changed significantly in terms of their difficulty level, we wondered whether the items rejected by Mantel-Haenszel statistic could be combined with the successfully translated items in order to still form a unidimensional scale. As noted, the items “I am depressed” and “I am very depressed” differ in terms of their difficulty, not in terms of their content: as both items are focused on depression. Accordingly, we Rasch analyzed the Filipino GDS data without the English-language data and inspected its fit statistics. Infit and outfit statistics in the translated items ranged from .55 to 1.40. This is a slightly better range than the English language (values of 1.0 are best), indicating that the thirty translated items indeed form a unidimensional measure of depression. Acceptable fit statistics support the conclusion that the translated items changed only in difficulty, not in quality. That is, the meaning of the items still indicates depression, not some other construct.

Numerous short forms of the GDS have been translated in various languages, including Chinese, Spanish, Iranian, and Malay (e.g., 13, 16, 17, 33). The current article reported a translation of the GDS into Filipino. Because the GDS assesses mainly depressive mood and its cognitive symptoms, such translations may be seen as an effort to establish a ruler for depressive mood that spans various cultures. As shown by our results, the Filipino long form of the GDS cannot be used as a simple replacement for long form of the English GDS. Twelve of the 30 items were significantly different in terms of the difficulty by the Mantel-Haenszel statistic. These items would need to be retranslated and piloted again in another sample of participants with depression, where the difficulty of the English- and Filipino-versions of each item would again be compared.

Nevertheless, it is also true that the 30 Filipino items form a unidimensional measure of depressive mood, as shown by their acceptable fit statistics. As such, the content has not changed, only the difficulty value of 12 of its items. This supports the conclusion that the 30 Filipino translated items can be used in research as a valid, interval-level measure of depression.

We also wondered whether the 18 successfully translated items could be fashioned into a short form that obeyed the requirements of the Rasch model, while conforming to the content of depression as given in the DSM-5. Table 2 compares the items of the GDS long and short forms to the DSM-5 diagnostic criteria for Depressive Episode.

A Short Form of the GDS: The Filipino GDS-15

As can be seen, both the GDS long and short forms primarily emphasize depressive mood. Neither form asks about weight loss or weight gain, insomnia or hypersomnia, or recurrent thoughts of death (i.e., all of which were omitted from the original GDS because they could be present in older adults in the absence of

depression). Two questions (question 13 and question 18) ask about worry. One item (item 6) asks about “thoughts you can’t get out of your head,” without specifying its content. These “thoughts” could be recurrent thoughts of death (criterion 9), but also represent obsessions. Consequently, item 6 was considered an anxiety-related item, not part of the DSM criteria for depressive episode. The GDS-15 drops all the anxiety items, and also drops question 11, which is related to psychomotor issues (criterion 5). The number of items in depressive mood (criterion 1) and loss of interest (criterion 2) are also reduced.

The last column of Table 2 compares the content of the successfully translated Filipino items to the GDS and DSM diagnostic criteria. The content coverage is similar between the Filipino GDS and the English GDS-15. The English GDS 15 has 7 depressive mood items, while the surviving Filipino GDS items have 6. Both have three items oriented to loss of interest or pleasure. Additionally, the Filipino items include one related to psychomotor issues, a theme which does not appear in the GDS-15. The three anxiety items were successfully translated, but do not appear in the GDS-15. Many depression scales contain anxiety items, and their inclusion in the GDS long form and omission from the GDS-15 suggests that Sheikh and Yesavage [19] regarded anxiety as related to depressive symptoms, but not core to depressive symptoms.

As a final step, then, we decided to Rasch analyze 15 of the 18 successfully translated items, omitting the 3 items that assessed anxiety. Again, our results indicated that all items satisfy the Rasch model, thereby creating what we term the Filipino GDS-15. Infit and outfit statistics ranged from a low of .66 to a high of 1.45, a range which is acceptable for dichotomous data. Internal consistency was more than adequate at .79. These statistics support the conclusion that the Filipino GDS-15 can be used as a short measure of depression that provides a genuine interval level of measurement.

However, we emphasize that the the Filipino GDS-15 is not a “drop in replacement” for the English GDS-15. The two forms have only 5 items in common. There are a variety of short forms of the GDS in circulation, and as always, researchers are advised to understand how the particular form available to them might differ from the original GDS and the GDS-15 published by Sheikh and Yesavage [19], as these provide the two most important reference points in the “GDS tradition.”

References

1. Philippine Statistic Authority. Highlights of the Philippine Population 2015 Census of population. <https://psa.gov.ph/statistics/census/population-and-housing>.
2. Sadock BJ, Sadock VA. Synopsis of Psychiatry. Philadelphia: Lippincott Williams & Wilkins. 10th ed. 2007; 533-544.
3. Taqui AM, Itrat A, Qidwa W, et al. Depression in the Elderly: “Does Family System Play a Role?” A Cross Sectional Study. BMC Psychiatry. 2007; 7: 57.
4. Lebowitz BD, Pearson JL, Schneider, et al. Diagnosis and Treatment of Depression in Late Life: Consensus Statement

- Update. *JAMA*. 1997; 278: 1186-1190.
5. Maletta G, Mattox KM, Dysken M. Guidelines for Prescribing Psychoactive Drugs. *Geriatrics*. 2000; 55: 65-79.
 6. Nierenberg AA. Current Perspectives on the Diagnosis and Treatment of Major Depressive Disorders. *American Journal of Management Care*. 2001; 7: 353-366.
 7. Heok KE, Ho R. The Many Faces of Geriatric Depression. *Current Opinion in Psychiatry*. 2008; 21: 540-545.
 8. Pepersack T, Breucker SD, Mekongo YPN, et al. Correlates of Unrecognized Depression Among Hospitalized Geriatric Patients. *Journal of Psychiatric Practice*. 2006; 12: 160-167.
 9. Kulathunga M, Umayal S, Somaratne S, et al. Validation of the Geriatric Depression Scale for an elderly Sri Lankan clinic population. *Indian Journal of Psychiatry*. 2010; 52: 254-256.
 10. Yesavage JA., Brink TL, Rose TL, et al. Development and Validation of a Geriatric Depression Screening Scale: A preliminary report. *Journal of Psychiatric Research*. 1983; 17: 37-49.
 11. McDowell I. *Measuring Health*. New York: Oxford University Press. 2006; 342.
 12. Kieffer KM, Reese RJ. A Reliability Generalization Study of the Geriatric Depression Scale. *Educational and Psychological Measurement*. 2002; 62: 969-994.
 13. The EE. Validation of Malay Version of Geriatric Depression Among Elderly Inpatient. <http://www.priory.com/psych/MalayGDS.htm>
 14. Gautam R, Houde SC. Geriatric Depression Scale for Community-dwelling Older Adults in Nepal. *Asian Journal of Gerontology & Geriatrics*. 2011; 6: 93-99.
 15. Kim JY, Park JH, Lee JJ, et al. Standardization of the Korean Version of the Geriatric Depression Scale: Reliability, Validity, and Factor Structure. *Psychiatry Investigation*. 2008; 5: 232-238.
 16. Malakouti SK, Fatollahi P, Mirabzadeh A, et al. Reliability, Validity and Factor Structure of the GDS-15 in Iranian elderly. *International Journal of Geriatric Psychiatry*. 2006; 21: 588-593.
 17. Cheng ST, Chan AC. A Brief Version of the Geriatric Depression Scale for the Chinese. *Psychological Assessment*. 2004; 16: 182-186.
 18. Melgar MIE, Salum D, Barzaga A, et al. Psychometric Properties of the Filipino version of the Geriatric Depression Scale. *Journal of Humanities and Social Sciences*. 2010; 15: 125-129.
 19. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): Recent Evidence and Development of a Shorter Version. In McDowell, I, ed. *Measuring Health*. New York: Oxford University Press 2006; 359-364.
 20. De Craen AJ, Heeren TJ, Gussekloo J. Accuracy of the 15-item Geriatric Depression Scale (GDS-15) in a Community Sample of the Oldest Old. *International Journal of Geriatric Psychiatry*. 2003; 18: 63-66.
 21. Chiang KS, Green EK, Cox EO. Rasch Analysis of the Geriatric Depression Scale- Short Form. *The Gerontologist*. 2009; 49: 262-275.
 22. Leshner EL, Berryhill JS. Validation of the Geriatric Depression Scale-Short Form Among Inpatients. *Journal of Clinical Psychology*. 1994; 50: 256-260.
 23. Alden D, Austin C, Sturgeon R. A Correlation between the Geriatric Depression Scale Long and Short Forms. *Journal of Gerontology*. 1989; 44: 37-43.
 24. Conradsson M, Rosendahl M, Littbrand H, et al. Usefulness of the Geriatric Depression Scale 15-item Version Among Very Old People With and Without Cognitive Impairment. *Aging & Mental Health*. 2013; 17: 638-645.
 25. Sugishita K, Sugishita M, Hemmi I, et al. A Validity and Reliability Study of the Japanese Version of the Geriatric Depression Scale 15 (GDS-15- J). *Clinical Gerontologist*. 2016; 40: 233-240.
 26. Paradela EM, Lourenco RA, Veras RP. Validation of Geriatric Depression Scale in a General Outpatient Clinic. *Revista de Saude Publica*. 2005; 39: 918-923.
 27. Reuland DS, Cherrington A, Watkins GS, et al. Diagnostic Accuracy of Spanish Language Depression-Screening Instruments. *The Annals of Family Medicine*. 2009; 7: 455-462.
 28. Nyunt MSZ, Fones C, Niti M, et al. Criterion-based Validity and Reliability of the Geriatric Depression Scale (GDS-15) in a Large Validation Sample of Community-living Asian Older Adults. *Aging and Mental Health*. 2009; 13: 376-382.
 29. Chachamovich E, Fleck MP, Power M. Is Geriatric Depression Scale-15 a Suitable Instrument for Measuring Depression in Brazil? Results of a Rasch Analysis. *Psychology, Health & Medicine*. 2010; 15: 596-606.
 30. Wai Kwong T, Wong E, Chiu HK, et al. The Geriatric Depression Scale Should be Shortened: Results of Rasch Analysis. *International Journal of Geriatric Psychiatry*. 2005; 20: 783-789.
 31. Cohen RJ, Swerdlick ME. *Psychological Testing and Assessment: An Introduction to Tests and Measurement*. NY: McGraw Hill. 2005; 141-174.
 32. Mantel N, Haenszel W. Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease. *Journal of The National Cancer Institute*. 1959; 22: 719-748.
 33. Fernández-San Martín MI, Andrade C, Molina J, et al. Validation of the Spanish Version of the Geriatric Depression Scale (GDS) in primary care. *International Journal of Geriatric Psychiatry*. 2002; 17: 279-287.