Assessment of depression in a geriatric inpatient cohort: A comparison of the BDI and GDS

Bret G. Bentz (Texas Woman’s University, USA) and James R. Hall (University of North Texas, USA)

(Received October 31, 2006/ Recibido 31 de octubre 2006)
(Accepted March 1, 2007/ Aceptado 1 de marzo 2007)

ABSTRACT. It was the purpose of this investigation to directly compare the Beck Depression Inventory (BDI) and Geriatric Depression Scale (GDS) in their ability to assess and classify depression in a geriatric inpatient population. A retrospective chart review of 158 consecutively admitted patient’s medical records to a geriatric inpatient unit was conducted. Data collected from chart reviews included diagnoses, demographic information, scores on both the BDI and GDS, and scores from several neuropsychological tests. Items from both the BDI and GDS were factor analyzed to obtain orthogonal constructs. In addition, the items of both the BDI and GDS were entered into a discriminant function analysis to investigate their ability to classify depression diagnosis. Finally, the factor scores and total scores from each questionnaire were correlated with measures of neuropsychological function. The results indicated that both the BDI and GDS have multi-factor structures and have questionable utility in the classification of depression. Furthermore, the Vegetative Symptoms factor of the BDI was found to significantly correlate with several measures of neuropsychological function. The results were discussed in terms of the relative clinical utility of these two self-report depression measures.


RESUMEN. El propósito de esta investigación fue comparar la habilidad para evaluar y clasificar la depresión del Beck Depression Inventory (BDI) y de la Geriatric Depression Scale (GDS) en su capacidad para evaluar y clasificar la depresión en una cohorte de pacientes geriátricos ingresados. Se realizó una revisión retroactiva de 158 expedientes médicos consecutivamente ingresados a la unidad de reingreso geriátrico. Se recogieron datos de las revisiones de los expedientes, incluyendo diagnósticos, información demográfica, puntuaciones en el cuestionario Beck Depression Inventory (BDI) y Geriatric Depression Scale (GDS), y puntuaciones de varios pruebas neuropsicológicas. Los ítems de ambos cuestionarios fueron factorizados para obtener constructos ortogonales. Además, los ítems de ambos cuestionarios BDI y GDS fueron introducidos en un análisis de función discriminante para investigar su capacidad para clasificar la enfermedad depresiva. Finalmente, las puntuaciones de los factores y las puntuaciones totales de cada cuestionario fueron correlacionadas con medidas de función neuropsicológica. Los resultados indicaron que ambos cuestionarios tienen estructuras multifactoriales y tienen cuestionable utilidad en la clasificación de la depresión. Además, el factor de síntomas vegetativos del BDI se encontró correlacionado significativamente con varias medidas de función neuropsicológica. Los resultados se discutieron en términos de la utilidad clínica relativa de estas dos medidas de depresión auto-reportadas.

1 Correspondence: Texas Woman’s University. Department of Psychology and Philosophy. PO Box 425470. Denton, TX 76204 (USA). E-Mail: bbentz@mail.twu.edu.
A number of self-report measures have been constructed to assess depression in a wide variety of populations. Two widely used self-report measures of depression with the geriatric population include the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, and Erbaugh, 1961) and the Geriatric Depression Scale (GDS; Yesavage et al., 1982). The BDI was developed to assess a range of cognitive and somatic symptoms and is widely used in research and clinical practice. The BDI has been confirmed as both reliable and valid for use in normal and psychiatric populations (Beck, Steer, and Garbin, 1988). Although not designed specifically for geriatric populations, the BDI has been reported to have good reliability with older adults (Gallagher, Nies, and Thompson, 1982) and to measure depressive symptoms independent of age (Steer, Rissmiller, and Beck, 2000). Numerous studies have investigated the factor structure of the BDI (Endler, Rutherford, and Denisoff, 1999; Golin and Hartz, 1979; Killgore, 1999; Schotte, Maes, Cluydts, De Doncker, and Cosyns, 1997; Shafer, 2006) and have found a two factor solution including a cognitive symptom factor and a somatic symptom factor. Still other investigations have reported a range of between three and seven factors (Beck and Lester, 1973; Gibson and Becker, 1973; Louks, Hayne, and Smith, 1989) depending on the extraction procedure and population used in the study. It is important to note, however, that the factor structure of the BDI using a geriatric population has yet to be investigated. Thus, the factor analytic investigation of the BDI reveals a depression measure that is multi-factorial and is lacking in evidence of its structure in a geriatric cohort.

Perhaps the most empirically difficult requirement of a self-report measure of depression in an elderly population is the measurement of depression independently of cognitive function. Stated simply, cognitive decline associated with age, Alzheimer’s disease, or stroke can adversely impact the ability of self-report measures to accurately assess depression. Therefore, it is advantageous for a self-report instrument to measure...
depression independently of cognitive function. The utility of the BDI in evaluating depression in the elderly with cognitive dysfunction has been questioned due to the effect of somatic type questions on the assessment of medically ill patients (Wagle, Ho, Wagle, and Berrios, 2000). Used as a screening measure, Wagle et al. (2000) noted that the BDI showed high rates of false negatives and concluded that the instrument was not an ideal measure of depression in a group of Alzheimer’s patients. The specific relationship of the BDI and cognitive dysfunction in this study was not established. This relationship may be both clinically and empirically useful due to the possible impact that cognitive function may have on the measurement of depressive symptoms and future construction of a depression measure.

Another widely used self-report measure of depression, the Geriatric Depression Scale (GDS; Yesavage et al., 1982), was developed to assess the cognitive, emotional and behavioral symptoms of depression specifically for the elderly population. The GDS has been validated as a useful measure of depression in older populations (Abraham, Wofford, Lichtenberg, and Holroyd, 1994; Montorio and Izal, 1996; Sheik, Yesavage, Brooks, and Friedman, 1991). Research on the factor structure of the GDS has revealed from one to five factors depending on the population under study (Adams, Matto, and Sanders, 2004; Salamero and Marcos, 1992; Sheikh et al., 1991).

In similar findings to the BDI, an association between the GDS and cognitive functioning has been suggested in the literature. Lichtenberg, Ross, Millis, and Manning (1995) found the GDS to predict level of cognitive function in a sample of geriatric medical patients. The GDS was found to predict scores of dementia and memory function suggesting that the measurement of depression is adversely impacted by cognitive decline. Conversely, Feher, Larrabee, and Crook (1992) concluded that the GDS was a valid measure of depression in a sample of outpatients with dementia. Finally, several other investigations (Gilley and Wilson, 1997; Rubin, Veiel, Kinscherf, Morris, and Storandt, 2001) have suggested that the GDS has limited value in the assessment of depression in Alzheimer’s patients. The limitation was reportedly due to the association between GDS scores and cognitive functioning (Gilley and Wilson, 1997).

Thus, a controversy exists as to the ability of self-report depression measures to accurately assess depressive symptoms independent of cognitive function in a geriatric population. It was the goal of the present descriptive ex post facto study (Montero and León, 2007; Ramos-Álvarez, Valdés-Conroy, and Catena, 2006) to directly compare the BDI and the GDS in their ability to assess depression in a geriatric population. The factor structure and ability to classify depression diagnosis were analyzed. In addition, the relationships between the specific constructs being measured by each depression scale and level of cognitive functioning were directly compared.

Method

Subjects

A retrospective chart review of consecutively admitted patient’s medical records to a geriatric psychiatry inpatient unit was conducted at the Flow Rehabilitation Hospital in Denton, Texas. All patients were initially evaluated upon admission by the attending
psychiatrist and psychiatric diagnoses were assigned according to DSM-IV criteria (American Psychiatric Association, 1994). Psychiatric diagnosis was made independently of the psychological testing. Participants were administered the CERAD, BDI, and GDS separately from the admission procedures.

Criteria for inclusion in the study were medical records that included data of an age greater than 64 years and completion of psychological testing that included both the BDI and the GDS self-report depression measures, plus several neuropsychological tests. All psychological testing was completed within five days of admission. A total of 158 patients (N = 158) were identified for inclusion in the study. The average age of the participants was 78.3 years (SD = 7), ranging from 65 to 101 years. The distribution of gender within the sample included 120 female (75.9%) and 38 male (24.1%) participants. In addition, the sample was primarily Caucasian (n = 155, 98.1%), with the remainder of participants including African-American (n = 2, 1.3%) and Hispanic (n = 1, 0.6%) subjects.

The physician diagnoses of depression and dementia were recorded from the medical record and were made independently of the self-report depression and neuropsychological assessment measures obtained from the medical record. Any participant with an Axis I diagnosis of Major Depression, Dysthymic Disorder, or Depression NOS, were classified as a diagnosis of depression. In addition, any participant with an Axis I diagnosis of Dementia of the Alzheimer’s type, Dementia due to a general medical condition, and Dementia NOS were classified as a diagnosis of dementia. Of the total sample, 85 (53.8%) participants were found to have a diagnosis of both depression and dementia. Forty-four (27.8%) participants had a diagnosis of dementia only. Twenty-five (15.8%) participants had a diagnosis of depression only. Finally, 4 (2.5%) participants were found to have no diagnosis of a depressive disorder or dementia.

Instruments

- The Beck Depression Inventory (BDI; Beck et al., 1961) is a 21-item self-report questionnaire designed for the assessment of depressive symptomatology in the general population. Its factor structure and validity has been established in student (Killgore, 1999), non-clinical (Endler et al., 1999), and mildly depressed populations (Golin and Hartz, 1979). The self-report format of the BDI was modified by having an examiner read all items to the participants in order to minimize blank responses and enhance uniformity.

- The Geriatric Depression Scale (GDS; Yesavage et al., 1982) is a 30-item self-report questionnaire designed for the assessment of depressive symptomatology in a elderly population. Its factor structure, reliability, and validity has been established in older depressed populations (Abraham et al., 1994; Sheikh et al., 1991). The self-report format of the GDS was modified by having an examiner read all items to the participants in order to minimize blank responses and enhance uniformity.

- Neuropsychological measures: The neuropsychological tests from the Consortium to Establish a Registry for Alzheimer’s Disease assessment battery (CERAD; Morris et al., 1989) were used in the screening of all participants. The reliability
and validity of the CERAD in the assessment of dementia has been established (Morris et al., 1989). These tests included measures of general cognitive dysfunction (Mini-Mental state), verbal fluency, confrontational naming, constructional abilities, immediate verbal memory, delayed verbal memory, and verbal memory recognition. In addition, the difference between delayed memory and recognition memory was calculated by subtracting the delayed verbal memory score from the recognition score. This recognition minus recall measure was included because of its reported importance in the assessment of Alzheimer’s disease and differential diagnosis (Caine, 1981). Finally, the Extended Orientation Exam from the Wechsler Memory Scale was included as a general measure of orientation (Wechsler, 1945).

Results

Factor analyses

In order to extract meaningful constructs as measured by both the BDI and the GDS in a geriatric inpatient population, factor analyses were conducted on the items of each measure. Principal component analyses with varimax rotation were utilized to extract orthogonal constructs. Only factors that achieved an eigenvalue greater than one and had three or greater items loading on the factor were included as meaningful constructs.

BDI analysis: The individual 21-item responses of the BDI from all participants included in the study were used in the analysis. The principle components factor analysis on the BDI found a total of five factors with an eigenvalue greater than one. However, two factors had item loading results that included less than three items and, therefore, were not included as meaningful constructs in the remaining analysis. The three factor solution explained a total of 44% of the common variance. The three meaningful BDI factors and item loadings are presented in Table 1. Factor 1 was named Negative Self-Image, Factor 2 Anhedonia, and finally Factor 3 was named Vegetative Symptoms.

| TABLE 1. Factor structure and item loadings for the BDI. |
|-----------------|----------------|
| **Factor 1: Negative Self-Image** | **Loading** |
| I feel irritated all of the time now. | .71 |
| I can’t make decisions at all anymore. | .64 |
| I feel I am being punished. | .64 |
| I can’t do any work at all. | .63 |
| I blame myself for everything bad that happens. | .60 |
| I hate myself. | .57 |
| I am so worried about my physical problems that I cannot think about anything else. | .51 |
| I feel guilty all of the time. | .48 |
| I believe that I look ugly. | .48 |
The individual 30-item responses of the GDS from all participants included in the study were used in the analysis. The principle components factor analysis on the GDS found a total of ten factors with an eigenvalue greater than one. However, five factors had item loading results that included less than three items and, therefore, were not included as meaningful constructs in the remaining analysis. The five factor solution explained a total of 44.4% of the common variance. The five meaningful GDS factors and item loadings are presented in Table 2. The depression scale items are presented in terms of the actual questions on the GDS. However, the item loadings for the reverse scored questions are noted and reflect the reverse score. Factor 1 was named Hopelessness, Factor 2 Social Isolation, Factor 3 Negative Affect, Factor 4 Irritability, and finally Factor 5 was named Worry.

### TABLE 1. Factor structure and item loadings for the BDI. *(Continued)*

<table>
<thead>
<tr>
<th>Factor 2: Anhedonia</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel that the future is hopeless and that things cannot improve.</td>
<td>.81</td>
</tr>
<tr>
<td>I am dissatisfied or bored with everything.</td>
<td>.63</td>
</tr>
<tr>
<td>I feel I am a complete failure as a person.</td>
<td>.61</td>
</tr>
<tr>
<td>I am so sad or unhappy that I can’t stand it.</td>
<td>.57</td>
</tr>
<tr>
<td>I have lost all of my interest in other people.</td>
<td>.57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 3: Vegetative Symptoms</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have no appetite at all anymore.</td>
<td>.80</td>
</tr>
<tr>
<td>I have lost more than 15 pounds.</td>
<td>.67</td>
</tr>
<tr>
<td>I wake up several hours earlier than usual and find it hard to get back to sleep.</td>
<td>.55</td>
</tr>
</tbody>
</table>

### TABLE 2. Factor structure and item loadings for the GDS.

<table>
<thead>
<tr>
<th>Factor 1: Hopelessness</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you think that it is wonderful to be alive now?</td>
<td>.78</td>
</tr>
<tr>
<td>Are you hopeful about the future?</td>
<td>.73</td>
</tr>
<tr>
<td>Do you feel that your situation is hopeless?</td>
<td>.59</td>
</tr>
<tr>
<td>Do you feel happy most of the time?</td>
<td>.55</td>
</tr>
<tr>
<td>Do you find life very exciting?</td>
<td>.44</td>
</tr>
<tr>
<td>Do you enjoy getting up in the morning?</td>
<td>.43</td>
</tr>
<tr>
<td>Are you in good spirits most of the time?</td>
<td>.41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 2: Social Isolation</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you prefer to stay at home rather than going out and doing new things?</td>
<td>.76</td>
</tr>
<tr>
<td>Do you prefer to avoid social gatherings?</td>
<td>.74</td>
</tr>
<tr>
<td>Is it hard for you to get started on new projects?</td>
<td>.55</td>
</tr>
</tbody>
</table>
In order to compare the ability of both the BDI and the GDS to classify depression diagnosis, two discriminant function analyses were conducted. The individual item ratings from both the BDI and the GDS were entered into two separate analyses and used to predict depression diagnosis.

**BDI analysis:** All 21-items of the BDI were used in a discriminant function to predict depression diagnosis. The function was found to significantly predict diagnosis ($\chi^2_{21} = 52.2$, $p < .001$). A summary of the classification results is presented in Table 3. A total of 74.1% of the originally grouped cases were correctly classified. Specificity of the BDI was found to be 80.9%. However, the sensitivity was slightly lower at 72.5%.

**GDS analysis:** All 30-items of the GDS were used in a discriminant function to predict depression diagnosis. The function was found to significantly predict diagnosis ($\chi^2_{30} = 68.19$, $p < .001$). A summary of the classification results is presented in Table 3.

<table>
<thead>
<tr>
<th>Original depression diagnosis</th>
<th>Predicted group membership</th>
<th>% correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>79</td>
<td>50</td>
</tr>
<tr>
<td>No</td>
<td>9</td>
<td>24.1</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>74.1</td>
</tr>
</tbody>
</table>
4. A total of 79.2% of the originally grouped cases were correctly classified. Specificity of the GDS was found to be 82.6%. Sensitivity was 81.3%.

**TABLE 4.** Classification results of the GDS discriminant function analysis.

<table>
<thead>
<tr>
<th>Original depression diagnosis</th>
<th>Predicted group membership</th>
<th>% correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>87</td>
<td>55.1</td>
</tr>
<tr>
<td>No</td>
<td>8</td>
<td>24.1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>79.2</td>
</tr>
</tbody>
</table>

**Correlational analyses**

In order to investigate the relationship between level of cognitive functioning and its impact on the two measures of depression, two correlational analyses were conducted. Specifically, the total scores and factor scores obtained from both the BDI and the GDS were correlated with the several neuropsychological test scores. In each analysis, Type I error was controlled with a Bonferroni correction of the alpha level to \( p < .001 \) due to the multiple statistical tests being completed.

**BDI analysis:** Table 5 presents the correlation results of the BDI scores and the neuropsychological measures. After the Bonferroni correction of the alpha level, the Vegetative Symptoms factor of the BDI was found to significantly correlate with the Mini-Mental State, Word List Memory, Word List Recognition, Recognition Minus Recall, and Orientation scores. No other significant correlations were found between the neuropsychological measures and the factor or total scores of the BDI.

**GDS analysis:** Table 6 presents the correlation results of the GDS scores and the neuropsychological measures. After the Bonferroni correction of the alpha level, no
significant correlations were found between the neuropsychological measures and the factor or total scores of the GDS.

TABLE 6. Correlation matrix of the GDS and neuropsychological measures.

<table>
<thead>
<tr>
<th>Neuropsychological measures</th>
<th>Geriatric Depression Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td>.06</td>
</tr>
<tr>
<td>Modified Boston Naming</td>
<td>.02</td>
</tr>
<tr>
<td>Mini-Mental State Exam</td>
<td>.03</td>
</tr>
<tr>
<td>Word List Memory</td>
<td>.10</td>
</tr>
<tr>
<td>Constructional Praxis</td>
<td>-.09</td>
</tr>
<tr>
<td>Word List Recall</td>
<td>.07</td>
</tr>
<tr>
<td>Word List Recognition</td>
<td>.03</td>
</tr>
<tr>
<td>Word List Recognition Minus Recall</td>
<td>-.02</td>
</tr>
<tr>
<td>Extended Orientation Exam</td>
<td>-.01</td>
</tr>
</tbody>
</table>

Note. F1: Hopelessness; F2: Social Isolation; F3: Negative Affect; F4: Irritability; F5: Worry.

Discussion

The purpose of this study was to directly compare the Beck Depression Inventory and the Geriatric Depression Scale in their ability to accurately assess depressive symptoms in an inpatient geriatric population. Each depression measure was factor analyzed, evaluated for ability to classify diagnosis, and correlated with several measures of neurocognitive functioning.

Both the BDI and the GDS were found to have multi-factorial factor structures in a geriatric inpatient population. Specifically, the BDI was found to have three orthogonal factors that included the constructs of Negative Self-Image, Anhedonia, and Vegetative Symptoms. This finding is contrary to the factor structure of the BDI reported by previous studies (Endler et al., 1999; Kilgore, 1999; Schotte et al., 1997) in which a structure including only two factors were found. However, it is likely that the observed difference in factor structure of the BDI is due to the subject population in the present study. Previous studies used various adult populations which suggests that there may be a difference in the constructs being measured by the BDI in a geriatric inpatient cohort. The GDS was found to have five orthogonal factors that included the constructs of Hopelessness, Social Isolation, Negative Affect, Irritability, and Worry. Previous research supports this finding of a multi-factorial structure to the GDS (Salamero and Marcos, 1992; Sheikh et al., 1991).

In the classification of depression diagnosis, the performance of the GDS was superior to the BDI. The percent of originally grouped cases correctly classified and the sensitivity of the GDS was found to be greater in comparison to the BDI. Only the specificity of the BDI was found to approach the performance of the GDS. Thus, in a geriatric inpatient population the GDS appears to classify depression more accurately,
especially with respect to a lower rate of false negatives. This finding supports the conclusions of several studies (Cabañero-Martínez, Cabrero-García, Richart-Martínez, Muñoz-Mendoza, and Reig-Ferrer, 2007; Lichtenberg, 1994).

However, the overall classification performance of both the BDI and GDS were marginal at best. Given that these self-report instruments are likely to be used as screening measures, false positive classification errors would be considered more tolerable than false negative errors. However, over one in four depressed participants for the BDI and almost one in five for the GDS were incorrectly classified as non-depressed. This high rate of false negative classification error indicates that the utility of both the BDI and GDS as depression screening measures in a geriatric population is questionable.

Finally, the correlational analyses of both the BDI and the GDS were completed with the results from the neuropsychological tests. No significant correlations between the factor and total scores of the GDS and the neuropsychological measures were found suggesting that the GDS is measuring depressive symptoms independently of cognitive function. However, several significant correlations between the Vegetative Symptoms factor of the BDI and the neuropsychological measures were found. This result suggests that appetite and sleep symptoms of depression are directly related to cognitive functioning and may help to explain the conclusions of Wagle et al. (2000). Specifically, Wagle et al. (2000) concluded that the BDI is not a model instrument for the assessment of depression in a group of Alzheimer’s patients. The current findings support this conclusion and suggest that the vegetative type questions may, in part, account for the poor performance of the BDI in the assessment of depression in a geriatric population.

In comparison to the BDI, it is concluded that the GDS is a superior measure of depressive symptoms in a geriatric inpatient population. The ability of the GDS to correctly classify depression diagnosis was found to be higher than the BDI. This superior ability to classify depression diagnosis was mainly seen in lower rates of false negatives on the part of the GDS. However, even though the GDS performance was superior to the BDI, overall the GDS still showed a rate of false negatives that was unacceptably high.

In relation to the neuropsychological testing, the GDS appears to assess depressive symptoms independently of level of cognitive functioning. Specifically, the Vegetative Symptoms factor of the BDI was found to be associated with several measures of neuropsychological functioning including orientation, memory, and Mini-Mental State scores. The GDS does not include vegetative type questions that assess appetite and sleep symptoms of depression. This lack of vegetative type questions appears to increase its ability to measure depression independently of cognitive functioning.

There were certain limitations to the design of this study which interpretation of the results should recognize. First, this study was retrospective in design. Future investigations of the relationship between depression measurement and cognitive functioning should attempt to employ a longitudinal design. Second, diagnoses and psychological testing were not blinded. Third, secondary diagnoses such as psychosis, anxiety, and other medical diagnoses were not controlled. Fourth, the current study used the original version of the BDI which has been revised to the BDI-II (Beck, Steer, and Brown, 1996). However, the two versions of the BDI are similar in their construction.
Finally, the study sample included a level of psychiatric illness that was severe. A future study should include a wide range of illness from normal to severely ill to increase the ability to generalize the results to a broad range of the geriatric population.

References


