An Analysis of Beck Depression Inventory 2nd Edition (BDI-II)

Mignote Hailu Gebrie*
School of Nursing, University of Gondar, Ethiopia

*Corresponding author: Mignote Hailu Gebrie, School of Nursing, College of Medicine & Health Sciences, University of Gondar, Gondar, Ethiopia

Submission: May 23, 2018; Published: July 05, 2018

Abstract

Beck Depression Inventory 2nd Edition (BDI-II) is a popular measure intended to assess the existence and severity of symptoms of depression in consistent with in the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV). It has been utilized both in clinical and research purposes with high reliability and validity as it is evidenced by plentiful psychometric analyses of the measure reported in the literatures. It has also translated in several languages and used in a range of cultural groups, diverse populations and settings. This paper presented the concept and goal of the measure and available psychometric analysis. Furthermore, gaps were identified in the literatures and provided recommendation about what type of research should be done in the future to address them.

Introduction

Current literatures revealed a wide-ranging evidence of the measurement of depression in both research and practice areas are necessary in understanding its extensive impact on individual’s health and quality of life. Depression has significant consequences in daily life and constitutes a major threat to chronic diseases [1]. The progressive decrements in function associated with many chronic medical illnesses may cause depression, and it is also associated with additive functional impairment [2]. As of other physical illnesses depression symptoms are highly prevalent in end stage renal disease (ESRD) patients yet it is still under-recognized and misdiagnosed [3,4]. It is often associated with adverse outcomes including mortality and non-adherence to medical treatment [3,5]. The Beck Depression Inventory (BDI) is a 21-item, self-report rating inventory that measures characteristic attitudes and symptoms of depression [6]. Beck Depression Inventory 2nd edition (BDI-II) is one of the most commonly used instruments in research and practice to measure the presence and severity of depression experienced during the past 2 weeks [7,8]. It is still a 21 item self-report inventory that reflect cognitive, affective and somatic components of depression, utilized in adolescents and adults [8-11]. Even though the original BDI relied upon the theoretical assumption that negativistic distorted cognitions would be the core characteristic of depression, the BDI-II does not reflect any particular theory of depression [12]. However, the BDI-II has been recently upgraded from BDI and the BDI-1A to make its symptom content more consistent with the diagnostic criteria used for major depressive disorders [13,14].

In the revision, the wording of all but three items has been changed from that of the BDI-1A for clarity [15]. The BDI-II contains four new symptoms, ‘agitation’, ‘worthlessness’, ‘concentration difficulty’ and ‘loss of energy’. The ‘weight losses, ‘body image change’ , ‘work difficulty’ and ‘somatic preoccupation’ symptoms from the BDI-1A were dropped [8,12,13]. Additionally, the 1-week time frame for rating each item was changed to 2 weeks and the response options for 14 old items were also revised [16]. Moreover, the items on appetite and sleep change were amended to evaluate the increase and decrease of these depression-related behaviors [12].

The scale is originally developed to identify the presence and severity of symptoms consistent with the criteria of the DSM-IV [17] but for diagnostic purposes, sleep pattern changes and appetite changes contain 7-point ratings to note increases or decreases in behavior [18]. Even though it has been criticized that the use of this scale in medically ill is challenging due to the somatic items included in the score, the BDI outnumbers the other measures in the amount of published research (there are more than 7,000 studies so far using this scale [13,19]. On the other hand, Arnau et al. [20] determined that the BDI-II yields reliable, internally consistent and valid scores in a primary care medical setting and concluded that the inclusion of somatic items was appropriate for their medical sample.

The BDI-II assesses 21 symptoms and attitudes which include Mood, Pessimism, Sense of Failure, Lack of Satisfaction, Guilt Feelings, Sense of Punishment, Self-dislike, Self-accusation, Suicidal
Wishes, Crying, Irritability, Social Withdrawal, Indecisiveness, Distortion of Body Image, Work Inhibition, Sleep Disturbance, Fatigability, Loss of Appetite, Weight Loss, Somatic Preoccupation, and Loss of Libido [21-23]. It is a widely used measure of depressive symptoms [24,25] and is considered gold standard for identifying depression in adults [11]. Despite its popularity, the factorial structure of the measure remains controversial [26]. The purpose of this paper is to provide an overview of the measure regarding its purpose, validity, reliability, criticisms and to identify gaps in the literature and recommend future researches.

**Analysis**

**Populations and settings**

The BDI-II is an immensely utilized screening instrument for depression, designed for use among individual adults and adolescents 13 years old and older [10,27]. The easy applicability and psychometric soundness of this scale have popularized its use in a wide range of populations and healthcare settings [19]. It is a reliable and valid measure of depression in a range of cultural groups and has been validated with both psychiatric and non-psychiatric populations in most of the countries including Africa [25]. Based on the current evidences the scale has been validated in psychiatric patients [28,29]. Moreover, several psychometric studies that validate the measure among diverse populations, settings and circumstances are published by a number of researchers in many languages and different cultures [1,21,30-32]. Populations for whom psychometrics are available for this measure include general populations of healthy adolescents and adults such as students [33] university students [34], college students [22,35] community adolescents [36] adult clinical inpatient [24], community-dwelling older and younger adults [27] and geriatric inpatients [14].

Psychometrics have also been studied in relation to various groups of patients having medical illnesses, including acute myocardial infarction [37], coronary heart disease [1], Human Immunodeficiency Virus [23], Cancer [38], Parkinsons disease [17], Chronic Kidney Disease [39] and End Stage Renal Disease [3] and persons with disabilities such as arthritis, spinal cord injury and amputation [18]. Even though the number of minority groups were too small and under-represented in the original study comprised of 91% White, 4% African American, 4% Asian American, and 1% Hispanic ethnic makeup [15], recent studies have indicated that the measure is valid and reliable for low-income African American sample of medical outpatients [8] and low income African American Suicide attempters [7] and the study of Adewuya et al. [40] also supports the use of the measure in Africa settings.

**Administration and Scoring**

The instructions of the BDI-II are straightforward and clearly stated. However, minimal training is required to administer or score the test and can be administered by paraprofessionals [18]. Although the BDI was initially designed to be administered by trained interviewers, it is most often self-administered; can be done with paper and pencil self-report in group or individual format self or oral administration and the data is collected retrospectively with individuals asked to report feelings consistent with their own over 2 weeks [13,18]. The time frame for ratings corresponds to that given by the DSM-IV for Major Depressive Disorder, and respondents are asked to describe themselves for the “Past Two Weeks, Including Today.” When self-administered, it takes 5-10 minutes and in oral administration it takes 15 minutes to complete [1,10,18].

The BDI-II is scored by summing the highest ratings for each of the 21 symptoms. Items are organized according to the severity of the content of alternative statements and each symptom is rated on a 4-point scale ranging from 0 (not) to 3 (severe) which covers cognitive, emotional/affective and somatic/vegetative symptoms with no sub scale and total scores can range from 0 to 63 [1,8,10,13,18,37]. Use the highest response when an item has greater than 1 severity rating. The scoring is criterion-referenced and performed by hand with scores 0-13 indicates minimal range, 14-19 mild depression, 20-28 moderate depression and 29-63 severe depression [10,12,15,18,24]. However, the interpretation of the final score requires a professional with clinical training and experience and no arbitrary cutoff score available for all purposes to classify different degrees of depression in this measure. Cutoffs have been recommended for specific medical populations. For instance, in post-myocardial infarction patients, the recommended cutoff value was greater than or equal to 16, with a sensitivity of 88.2% and a specificity of 92.1% [18].

Over time the BDI-II is translated and validated in numerous languages with high levels of reliability and validity across cultures though the original measure was developed in English. Among those, psychometric analysis is available for translation to Turkish, Japanese, Brazilian Portuguese, Persian, Indonesian, Swahili, Chinese, Spanish, Arabic, Swedish, Germany, Dutch and Icelandic respectively [1,10,21,27,30-34,36,41-43]. Analyses of BDI-II results have been conducted in a dichotomous variable as “depressed” versus “not depressed” and the severity has been rated as minimal or no depression, mild, moderate and severe. Adjustment of the Cutoff point can be made based on the characteristics of the sample and the purpose for use of the measure.

**Validity and reliability**

A number of studies have established the validity and reliability of BDI-II in different populations and settings. Regarding reliability of the measure, the internal consistency reliability was high on the original manual with a Cronbach’s α of 0.92 for the outpatient population and 0.93 for the college students, as reported by Smith & Erford [15] and Smarr & Keefer [18]. Similarly, the internal consistency reliability in a study on community-dwelling adults, depressed geriatric inpatients, HIV-positive adolescent, low income African American suicide attempters and low income African American sample revealed α = 0.90, 0.89, 0.80, 0.94, 0.90 respectively [7,8,10,14,44]. Furthermore, a good internal consistency reliability was reported in the scales translated version to Turkish (α = .90) [10], Japanese (α = .83) [31], Persian (α = .87) [32] and Indonesian (α = .90) [1].

Based on Smith and Erford [15] report about the original manual, the one-week test-retest correlation of the scale was 0.93,
in the Turkish adolescent it was stated as 0.89 [10] and a Persian version psychometric study demonstrated 0.73 [32] that shown a relative stability. The content validity of the scale has been improved by rewarding and adding items to assess DSM-IV criteria for depression. The measure has been evaluated for convergent and divergent validity in different studies and yields a positive correlation with Center for Epidemiologic Studies Depression Scale (r=0.69), Coolidge Axis II Inventory Depression sub scale (r=0.66), Coolidge Axis II Inventory Anxiety sub scale, Perceived Stress Scale (r=0.60) [28], Children’s Depression Inventory-II-Short [44], Hamilton Rating Scale for Depression (r = .66) [44] that supports the validity of the measure and negatively correlated with Short Psychological Well-Being Scale total score (r=-0.60) [10]. Besides, the two factors model of the measure was examined using confirmatory factor analysis and demonstrated it did not fit the data well in older adults with χ² (342, N=376)=1202.08, p<0.001, Comparative Fit Index (CFI)=0.69, root mean square error of approximation (RMSEA)=0.08 [10] where as it suggests an acceptable fit in low income African American suicide attempters (CFI=0.92 and IFI=0.93) [7].

Even though multiple exploratory and confirmatory factor analyses have been conducted on the measure that revealed a 2-factor or 3-factor structure solution still there is no clear evidence regarding which fits best. Therefore, additional large scale up-to-date factor analysis study is necessary to comprehend the measure well. Likewise, though the scale was cross validated in different cultures still there is a need of more research in this area to use it since cultural variation will have huge impact on the utilization as well as the interpretation of the measure.

Conclusion

BDI-II is validated, inexpensive, quick, and most frequently used self-rating scale to assess depression. It provides a fast, efficient way to assess depression in either a clinical or non-clinical environment and has several advantages for a researcher interested in assessing depressive symptoms [45-48]. These include simplicity in administration and scoring, excellent reliability and validity, availability of numerous translations, assessment of symptoms and time frame of measurement correspond to the DSM-IV criteria and usability for different populations and settings with low risk of harming the participants psychologically or in any other dimension of health that make it preferable for a number of context including chronically ill patients. On the other hand, disadvantages of the measure include the standardization samples are very small with no information regarding socioeconomic status, the need for health professionals for interpretation, is costly to buy and the challenge of inclusion of somatic symptoms in the measure in patients with medical conditions, although somatic symptoms have been shown to be an appropriate assessment for depression in primary care medical patients [20].

Psychometric studies of the BDI-II shown excellent internal consistency and one-week test-retest reliability on clinical samples and correlations with other tests maintaining to measure the construct of depression. However, there is a variation in determining the measure as two or three factor model as well as in the number of items per factor and which items are included on each factor. Even the two-factor solution that mostly used remains unclear as to what two-factor structure may perform best. Therefore, additional exploratory and confirmatory factor analysis studies must be undertaken to further understand the dimensionality underlying the measure.

Acknowledgement

I would like to extend my deepest appreciation to Dr. Cilia Willis for providing the avenue for me to do this analysis.

References


18. Smarr KL, Keefar AL (2011) Measures of depression and depressive symptoms: Beck depression inventory-II (BDI-II), center for epidemiologic studies depression scale (CES-D), geriatric depression scale (GDS), hospital anxiety and depression scale (HADS), and patient health questionnaire-9 (PHQ-9). Arthritis Care and Research 63(S11): S545-S546.


