 doi:10.20944/preprints201612.0083.v1

Essay

Depression Measurement Instruments: An Overview of the Top Depression Rating Scales

John H. Morgan

Dodge House, 415 Lincoln Way East, Mishawaka, IN 46544, USA;
faculty@gtfeducation.org; Tel.: +1-800-423-5983

ABSTRACT

The range of rating instruments in depression measurement and the depth of their analytical relevance constitutes a major development in this psychiatric and psychotherapeutic field of mental health. Though the competition is acute amongst these various instruments, the results for the public have been outstandingly positive. A depression rating scale is essentially a psychiatric measuring instrument utilized in the identification and ranking of depression severity within the patient. The scale provides the practitioner, psychiatrist or psychotherapist, with sufficient information to assess the severity of the depression plotted on the scale. Not used as a "diagnostic tool" itself, nevertheless, the depression rating scale does function as an effective device for designating and assigning a behavioral score which may, then, be used in establishing the severity of depression of value in the designation of a diagnosis and treatment formula. In this paper, we will take a close look at the leading depression rating scales and briefly summarize their scope of assessment value in rating depression.

KEY WORDS

depression; measurement scale; ratings; literature review

Both the range of rating instruments in depression measurement and the depth of their analytical relevance constitutes a major development in this psychiatric and psychotherapeutic field of mental health. Though the competition is acute amongst these various instruments, the results for the public have been outstandingly positive (Wechsler, Grosser, Busfield, 1963; Whooley, Avins, Miranda, Browner, 1997). A depression rating scale is essentially a psychiatric measuring instrument utilized in the identification and ranking of depression severity within the patient. The scale provides the practitioner, psychiatrist or psychotherapist, with sufficient information to assess the severity of the depression plotted on the scale. Not used as a "diagnostic tool" itself, nevertheless, the depression rating scale does function as an effective device for designating and assigning a behavioral score which may, then, be used in establishing the severity of depression of value in the designation of a diagnosis and treatment formula (Sharp, Lipsky, 2002). In this paper, we will take a close look at the leading depression rating scales and briefly summarize their scope of assessment value in rating depression.

Three types of scales are available, those that are completed by the researcher, those completed by the patient, and those requiring joint collaboration between therapist and patient.

 doi:10.20944/preprints201612.0083.v1

In this paper we will consider all three kinds. The Hamilton Depression Rating Scale, for example, is completed by the therapist and includes 21 questions offering 3 to 5 possible responses on the part of the patient, the higher the number the greater the depression. Labor intensive, the therapist (psychiatrist or psychotherapist) must interview the patient and rate, based on personal observation, the patient's symptoms. The HDRS is one of two most highly regarded measurement scales (Demyttenaere, 2003; Hamilton, 1960) with the Mongomery-Asberg Depression Rating Scale being the other. The MADRS has ten items requiring the therapist to rate the patient's response (Montogmery, Asberg, 1997). Another scale requiring the therapist to complete the report is the Raskin Depression Rating Scale (RDRS) which rates the patient's symptoms in three classification, viz., verbal reports, behavior, and secondary depression symptoms has been somewhat surpassed and is now under-used by the leading therapeutic clinics and researchers (Raskin, Schulterbrandt, Reatig, McKeon, 1969).

The verdict is out whether the patient-completed scale or the therapist-completed scale is better, the general consensus is that both have their relevance in diagnostic assessment. Of those completed by the patient, there is general agreement among researchers that the Beck Depression Inventory (BDI) is quite good with its 21-question self-report inventory covering a range of behavioral and emotional variables including irritability, fatigue, weight loss, lack of interest in sex, and feelings of guilt, hopelessness and fear of being punished (Beck, 1972). This instrument, predictably, is consistent with the DSM-IV diagnostic criteria (Yesavage, 1988). A self-administered scale which has gained much recognition for its thoroughness is the Geriatric Depression Scale (GDS) designed particularly for older patients and those suffering from signs of mild dementia. The simplicity of this scale is it use of "yes" and "no" responses rather than gradations (Katz, 19988; Zung, 1965) whereas the Zung Self-Rating Depression Scale (ZSRDS) emulates the GDS with preformatted answers. In this scale, 20 items divided between ten positively worded versus ten negatively worded questions which are rated on a scale of 1 to 4 based on four possible responses including "a little of the time," "some of the time," "good part of the time," and "most of the time" (Spitzer, Kroenke, Williams, 1999). A third measurement scale is the Patient Health Questionnaire (PHQ) consisting of several sets of scales, viz., the Patient Health Questionnaire-9 (PHQ-9) based on a self-reported 9-question series which is a modified version of the Primary Care Evaluation of Mental Disorders (Gilbody, House, Sheldon, Gilbody, 2005). The Patient Health Questionnaire-2 (PHQ-2), on the other hand, is a shorter version of the PHQ-9 focusing just on the existence of depression and loss of interest in daily activities. The Primary Care Evaluation of Mental Disorders (PRIME-MD) attempted to employ both patient and therapist responses addressing 27 items with follow-up clinical interviews but owing to its time-consuming application, it has recently been discarded in deference to the PHQ.

Before we commence a summary of the leading measurement scales for depression, it should be pointed out that even though there has been a call within the therapeutic professions for a more extensive evaluative scale for the detection of depression, alas, studies have shown that to date there is no evidence that increasingly refined measurement scales do not actually improve detection rates or treatment outcomes (DeJonge, Mangano, Whooley, 2007) and this particularly applies to the Hamilton Rating Scale for Depression (Zimmerman, Chelminski, McGlinchey, Posternak, 2008). The PHQ-9 has, however, shown a more promising utility of accuracy in specialized application such as in subgroups of patients with particular health maladies such as coronary disease (Kroenke, Spitzer, Williams, 2001). *(Note: The measurement*

 doi:10.20944/preprints201612.0083.v1

scales reviewed here have been treated alphabetically and their listing in no way suggests order of preference or effectiveness.)

The Beck Depression Inventory (and its later revisions BDI-1A and BDI-II) is a 21-questionnaire consisting of a multiple-choice self-reported inventory created by Aaron T. Beck, MD, and recognized as one of the most widely used psychometric tests for measuring the various levels of depression, from mile to severe (Craven, Rodin, Littlefield, 1988).. It reflects an industry wide shift away from the psychodynamic perspective and towards a phenomenology of the patient's own reported thoughts and experiences. The most recent version, the BDI-II, covers ages 13 through old age and measures depression covering a wide range of emotional responses (Beck, 1972). The earlier revised version, BDI-IA, modified the original BDI with an easier response mechanism for the patient and increased the correlation of inventory items in the questionnaire (Ambrosini, Metz, Bianchi, Rabinovich, Undie, 1991). These improvements were in response to the American Psychiatric Association's critique of the 1996 BDI published in the *Diagnostic and Statistical Manual of Mental Disorders*.

Similar to the original BDI, the BDI-II consists of 21 questions scored on a 0 to 3 valuation scale, higher scores suggesting increasingly severe depression (Steele, 2006). With a Pearson r of .71 correlation with the Hamilton Depression Rating Scale, the effective usefulness of the BDI-II is unsurpassed in the field including an even higher one-week test-retest reliability rating of Person r = 0.93 (Beck, Steer, Brown, 1996). Furthermore, the internal consistency was rated at (a-0.91) (Beck, Steer, Ball, Ranieri, 1996). With the BDI, the intensity of depression was subject to a quantitative assessment rather than a psychodynamic theory of analysis and, because the BDI indicates the level of depression quantitatively, it has the capacity to monitor and track changes in the patient over time (Beck, Ward, Mendelson, 1961). By 1998, the BDI was used in over 2,000 empirical studies (Richter, Werner, Heerlein, Kraus, Sauer, 1998).

Nevertheless, the BDI is subject to the commonly evidenced problems of all self-reporting inventories, viz., the danger of exaggeration or minimization by the responding patient (Bowling, 2005). For example, patients with concomitant physical illnesses may have their scores artificially inflated if such physical maladies such as fatigue suggest the illness rather than depression (Moore, Moore, Shaw, 1998). Attempting to account for this danger, Beck and colleagues created the Beck Depression Inventory for Primary Care (BDI-PC) instrument for short-term screening of participants measuring seven items from the BDI-II considered independent of concomitant physical illnesses (Storch, Roberti, Roth, 2004). This modification of the BDI resulted in a binary outcome finding of either "not depressed" or "depressed" for cutoff scores of 4 or above (Steer, Cavalier, Leonard, Beck, 1999). And, apart from the intent of the BDI measurement scale, some health care providers actually use it to research a quick diagnosis of a patient (Hersen, Turner, Beidel, 2007). Though the BDI is copyrighted and a fee is required for its use, there is actually no verifiable evidence that the BDI-II is any more valid or reliable than a range of other measuring tools to be discussed in this paper (Zimmerman, Chelminski, McGlinchey, Posternak, 2008) including those in the public domain such as the Patient Health Questionnaire-Nine Item (PHQ-9) (Kroenke, Spitzer, Williams, 2001).

The Burns Depression Checklist (BDC) is another highly regarded rating scale for depression developed by David Burns, MD. He is on the voluntary faculty of the Stanford

 doi:10.20944/preprints201612.0083.v1

University School of Medicine and has served as a statistical consultant for Stanford's Center for Interdisciplinary Brain Sciences Research. In addition, he has been a Visiting Scholar ath the Harvard medical School and Acting Chief of Psychiatry at the Presbyteian/University of Pennsylvania Medical Center in Philadephia. The BDC version of 1984 consisted of 15 questions but the upgraded version of 1996 now has 25 questions in the survey. The context for the answers is "during the past week from today" and is scored from 0 to 4, ranging from "not at all" to "extremely" and in Burns research it replaced the 1980 version of the BDI (burns, 1999).

Similar to Beck's proliferation of sub-scale testing and measurement tools, Burns has also created shorter measurement scales for depression, suicidal urges, anxiety, anger, and relationship satisfaction and, interestingly applicable to psychotherapeutic practice, he has developed an assessment scale for measuring the quality of the therapeutic alliance and effectiveness in the therapist/patient relationship. Judged highly reliable (above 0.90), these tests are completed by the patient and scored in less than a minute by the researcher. An interesting feature of Burns' approach to the process is not only having the patient complete the survey privately in the waiting room, being observed by the therapist during the pre and post testing period, but also has the patient assess the therapists and rate them based on warmth, empathy, and helpfulness during the process.

The Edinburgh Postnatal Depression Scale (EPDS) was designed specifically to identify patients who are suffering from postpartum depression and consists of a 10-item questionnaire (Cox, Holden, Sagovsky, 1987). These carefully selected items correlate to recognized clinical depression symptoms within psychiatric and psychotherapeutic practice including feelings of guilt, sleep disturbance, low energy, anhedonia, and suicidal thoughts. The comprehensive assessment is based on the total scoring of the EPDS with indications of higher levels of depressive symptoms correlated to higher scores (Vivilaki, Dafermos, Kogevinas, Bitsios, Lionis, 2009). This scale is commonly used within 8 weeks of postpartum as well as frequently used to identify depressive signs during pregnancy itself (Bergink, Kooistra, Lambregtse-van, den Berg, Wijnen, Bunevicius, van Baar, Pop, 2011). The EPDS is a widely employed depression screening mechanism and is used in a variety of countries in various languages including German, Spanish, French, and Polish (Garcia-Esteve, Ascaso, Ojuel, Navarro, 2003; Bergant, Nguyen, Heim, Ulmer, Dapunt, 1998; Vivilaki, Dafermos, Kogevinas, Bitsios, Lionis, 2009).

The Geriatric Depression Scale (GDS) is one of the earliest and most respected of depression measurement tools within the health care field and was developed in 1982 by J. A. Yesavage and colleagues (Yesavage, Brink, Rose, et al, 1982-83). Using a 30-item assessment questionnaire self-administered by the elderly patient in an attempt to identify evidence and levels of depression, the GDS questions are answered "yes" or "no" rather than using the otherwise more commonly five-category response set as discussed earlier. The simplicity of the positive/negative response mode is particularly effective with ill and mildly cognitively impaired patients and this test is utilized particularly in comprehensive geriatric assessment protocols and the score is based on a cumulative numbered grid ranging from 0 to 9 with the higher score reflecting increasingly severe depression.

 doi:10.20944/preprints201612.0083.v1

As with the other scales being reviewed here, the GDS is not designed to be used as a single diagnostic tool in identifying clinical depression even though it has unquestionably strong evidence of reliability and validity measured against other diagnostic instruments. A shorter version of the GDS called the GDS-SF contains only 15 questions (Sheikh, Yesavage, 1986) and this scale is available in several languages other than English including Arabic, Spanish, German, French, and Russian and this shorter version has been assessed as essentially adequate as a substitute for the long version (Lesher, Berryhill, 1994). As with the Beck Depression Inventory, the GDS has been evaluated against the Hamilton Rating Scale for Depression (HRS-D) as well as the Zung Self-Rating Depression Scale (SDS) and found a 92% sensitivity and a 89% specificity rating when evaluated against these diagnostic tools (Yesavage, Brink, Rose, et al, 1982-83).

The Hamilton Rating Scale for Depression (HRSD), sometimes called the Hamilton Depression Rating Scale (HDRS) and abbreviated HAM-D, is based on a multiple item questionnaire and addresses depression indicators with special attention to the evaluation of recovery from depression (Hedlund, Viewig, 1979). Initially developed by Max Hamilton in 1960 (Hamilton, 1960), it has gone through numerous revisions in 1966 (Hamilton, 1966), 1967 (Hamilton, 1967), 1969 (Hamilton, 1969), and 1980 (Hamilton, 1980). The particular attraction and popularity of the HRSD is that its questions are designed to rank the severity of depressive symptoms as determined by mood, feelings of guilt, suicidal thoughts, insomnia, agitation or retardation, anxiety, weight loss, and somatic symptoms broadly defined.

Originally 17 items comprised the instrument (HRSD-17) with four additional questions added for particular clinical information. Scored on a 3 or 5 point scale and taking 20 minutes to complete by the patient, a score of 0 to 7 constitutes normal with scores of 20 or above indicating moderate, severe, or very severe depression with the additional four questions used for further clinical assessment such as whether diurnal variations or paranoid symptoms were present in the patient. Helpful was a structured interview guide provided by the test (Williams, 1989).

Once considered the flag ship of depression rating scales in clinical research, the HRSD has fallen into less hallowed ground owing to its apparent over emphasis upon insomnia rather than suicidal thoughts as a major indicator of depression (Bagby, Ryder, Schuller, Marshall, 2004). It is argued, now, that an antidepressant has the capacity to show statistical efficacy even if suicidal thought increase while sleep improves (Kriston, Von Wolff, 2011). Furthermore, it is argued by detractors to the primacy of the HRSD that an antidepressant's side effects may actually increase sexual and gastrointestinal symptoms while remaining less effective in the treatment of depression (Berrios, Bulbena, 1990). Still highly regarded and extensively utilized, the HRSD has come under greater scrutiny owing somewhat to the proliferation of newer and more carefully tested depression measurement scales.

The Hospital Anxiety and Depression Scale (HADS), a common instrument used in medical practice, both private and institutional, to determine and assess the presence and level of anxiety and depression of a patient, was developed in 1983 by Zigmond and Snaith (Zigmond, Snaith, 1983). Using a 14 item scale for gathering patient responses, 7 are directly related to anxiety and the other 7 to depression. In an attempt to avoid reliance upon various aspects of these two conditions, anxiety and depression, owing to their common similarity to somatic

 doi:10.20944/preprints201612.0083.v1

symptoms of illness such as fatigue, insomnia, and hypersomnia, Zigmond and Snaith created this measuring instrument to focus precisely and specifically upon detection and measurement of identifiable traits in anxiety and depression within patients suffering from physical health problems.

In the HADS, each of the 14 items in the questionnaire are scored from 0 to 3 thereby assuring that the respondent can score from 0 to 21 for either anxiety or depression. This ordinal data is therefore based on a cut-off point for either anxiety or depression (Bjelland et al, 2002) and through a carefully constructed review of a large number of these studies they have identified a cut-off point of 8/21 for both anxiety and depression. The results has shown that anxiety (HADS-A) has rendered a specificity of 0.78 with an accompanying sensitivity of 0.9 while for depression (HADS-D) there is a rendering of 0.79 specificity and 0.83 sensitivity. There is no consensus, however, within the medical community as to the overall utility of the HADS in general practice dealing with either anxiety or depression, preferring rather that the patient is best served when this instrument is used as an unidimensional measurement for specifically psychological distress, thereby rendering its comprehensive utility less valued.

The Kutcher Adolescent Depression Scale (KADS), unlike most measuring instruments treated in this review article, is focused exclusively upon the care and treatment of adolescent depression (Kutcher, 2003). Stan Kutcher, M.D., Professor of psychiatry at the Dalhousie University in Canada, has chosen to highlight the effectiveness of pharmacological treatments of depression and, consequently, owing to pharmacological intervention in the biogenic symptoms of the adolescent patient, he addresses the assessment and diagnostics of depression severity over time which enhances the effectiveness and accuracy of his rating scale both for researcher and clinical practitioners (Brooks, Kutcher, 2003).

The commonly used version of the KADS is the one using 11 items for identification of adolescent depression for it is particularly effective in monitoring and measuring results in the testing and treatment of adolescents receiving medical care for acute depressive disorders. The 11 items are constructed so as to employ both standard and colloquial terminology found among adolescents and the responses gained from the measuring instrument are scored on a simple 4 choice scale (Brooks, 2004). Ten questions in the instrument are specifically about the frequency of depression symptoms in which the patient rates on a 4 point scale using these levels of severity, i.e., “hardly ever,” “much of the time,” “most of the time,” “all the time,” and the last question, which relates to the severity of suicidal thoughts (Brooks, Kutcher, 2003). The test is scored on a range from 0 to 33 and, unlike other rating scales discussed herein, there is no avenue for the introduction of sub-clinical presentations of findings nor does the scale provide for a range of responses from mild, moderate, or severe symptoms. The higher the score, the greater the severity of depression. The KADS has been extensively used in the U.S. and Canada and is lauded as one of the most refined scales of depression measurement among adolescents.

The My Mood Monitor Screen (M3 Checklist), though far from the depth and breadth of those measuring scales discussed thus far, was designed by Pedersen and colleagues (Pedersen, 2013) to be a quick and easy to administered as a validated, self-rated, multi-dimensional mental health symptom checklist which is designed to screen and monitor changes in mood and anxiety symptoms. Based on a 27-item questionnaire, it assesses a patient’s risk for developing

 doi:10.20944/preprints201612.0083.v1

depression, anxiety disorders, posttraumatic stress disorders, and bipolar disorders as well as behavioral dysfunctions and impairments related to drug and alcohol use and abuse. Though new to the field, it has been recognized by the National Committee for Quality Assurance (NCQA) as of 2013 and is recognized as a pre-validated mental health tool. The research instrument itself is hi-tech and available through smartphone use on an app called "WhatsMyM3" and on iTunes and Android. It calculates the user's score and then, by request of the user, the information may be sent directly to a designated health professional (Healy, 2012). Confidence is high in the effective assessment of the M3 Checklist owing to it having been assessed and validated against the Mini International Neuropsychiatric Interview (MINI), a standard diagnostic tool by Gaynes et al at the University of North Carolina (Gaynes, DeVeaugh-Geiss, Weir, Gu, MacPherson, Schulberg, Culpepper, Rabinow, 2010).

The Major Depression Inventory (MDI), designed by the World Health Organization, and is a self-reporting mood questionnaire constructed by a research group of psychiatrists led by Professor Per Bech, a psychiatrist based at Frederiksborg General Hospital in Denmark (Bech, Rasmussen, Olsen, Abildgaard, 2001). Distinctly different from the Beck Depression Inventory (BDI), it is designed to generate an ICD-10 or DSM-IV diagnosis of clinical depression in addition to an estimation of symptom severity (Bech, Rasmussen, Olsen, Noerholm, Abildgaard, 2001; Olsen, Jensen, Noerholm, Martiny, Bech, 2003). Thanks to the World Health Organization, another distinctive and attractive feature of the MDI is its availability free of charge and it is downloadable from the internet with an attached instructional manual on scoring and general usage (Cunningham, et al, 2011). And, needless to say, it is available in a dozen languages as well (Fawzi, Fawzi, Abu-Hindi, 2012). The treatment measurement scale ranges from mild, moderate, severe, to major depression. The higher the score, the more severe the level of depression (Olsen, Jensen, Noerholm, Martiny, Beck, 2003).

The Montgomeery-Asberg Depression Rating Scale (MADRS) was designed by British and Swedish researchers in 1979 to supplement the Hamilton Rating Scale for Depression (HAMD) owing to a desire to make more sensitive the research response to changes brought about due to antidepressants and other psychopharmacological interventions beyond what the HAMD proposed to study (Montgomery, Asberg, 1979). It is a ten-item diagnostic questionnaire designed by psychiatrists to measure the severity of depression particularly in patients with mod disorders (Williams, Kobak, 2008). There is, however, a consistency of correlation in the findings of both the MADRS and the HAMD (Williams, Kobak, 2008). The questionnaire covers symptoms of 10 kinds including apparent sadness, reported sadness, inner tension, reduced sleep, reduced appetite, concentration difficulties, lassitude, inability to feel, pessimistic thoughts, and suicidal thoughts (Svanborg, Asberg, 2001). What is helpful is the existence of a commonly used self-rating version of this scale called the MADRS-S which is frequently used by clinical practitioners (Cunningham, et al 2011) and employs an overall score ranging from 0 to 27 points easily interpreted by a professional (Svanborg, Asberg, 2001).

For better than 40 years the medical community in association with the behavioral science research community has engaged in the systematic development of depression measurement rating scales and public health has been well served by this enterprise. The 10 depression rating scales reviewed in this brief essay constitute the leading instruments employed in the field of psychiatry and psychotherapeutic counseling. The continued refinement of these

 doi:10.20944/preprints201612.0083.v1

and developing scales constitutes a fundamentally important contribution to the treatment of mental health and the complementarity of these instruments continue to elevate the quality of care provided by the medical science community.

---

*Research funded by the Lilly-Templeton Memorial Research Grant (939LTC) from the Zycon Corporation.

END NOTES

Ambrosini PJ, Metz C, Bianchi MD, Rabinovich H, Undie A (January 1991). "Concurrent validity and psychometric properties of the Beck Depression Inventory in outpatient adolescents". Journal of the American Academy of Child and Adolescent Psychiatry. 30 (1): 51–7.

Bagby RM, Ryder AG, Schuller DR, Marshall MB (2004). "The Hamilton Depression Rating Scale: has the gold standard become a lead weight?". American Journal of Psychiatry. 161 (12): 2163–77.

Bech P, Rasmussen N-A, Olsen LR, Noerholm V, Abildgaard W. The sensitivity and specificity of the Major Depression Inventory, using the Present State Examination as the index of diagnostic validity. J Affect Disord 2001; 66: 159-164.

Beck AT, Ward C, Mendelson M (1961). "Beck Depression Inventory (BDI)". Arch Gen Psychiatry. 4 (6): 561–571.

Beck AT, Steer RA and Brown GK (1996) "Manual for the Beck Depression Inventory-II". San Antonio, TX: Psychological Corporation.

Beck AT (1972). Depression: Causes and treatment. Philadelphia: University of Pennsylvania Press. p. 333.

Beck AT, Steer RA, Ball R, Ranieri W (December 1996). "Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients". Journal of Personality Assessment. 67 (3): 588–97.

Bergink V, Kooistra L; Lambregtse-van, den Berg MP; Wijnen H; Bunevicius R; van Baar A; Pop V (Apr 2011). "Validation of the Edinburgh Depression Scale during pregnancy". J Psychosom Res. 70 (4): 385–9.

Berrios G E & Bulbena A (1990) The Hamilton Depression Scale and the Numerical Description of the Symptoms of Depression. In Bech P & Coppen A (ed) *The Hamilton Scales*, Heildeberg, Springer, pp80-92.

Bjelland, I; et al. (2002). "The validity of the Hospital Anxiety and Depression Scale. An updated literature review". Journal of Psychosomatic Research. 52 (2): 69–77.

Bowling A (September 2005). "Mode of questionnaire administration can have serious effects on data quality". Journal of public health (Oxford, England). 27 (3): 281–91.

Brooks, Sarah; Kutcher, Stan (2003). "Diagnosis and Measurement of Adolescent Depression: A Review of Commonly Utilized Instruments". Journal of Child and Adolescent Psychopharmacology. 13 (3): 341–76.

Brooks, Sarah (2004). "The Kutcher Adolescent Depression Scale (11-Item)" (PDF). Journal of Child and Adolescent Psychopharmacology News. 9 (5): 4–6.

Cox J.L.; Holden J.M.; Sagovsky R. (1987). "Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale". Br J Psychiatry. 150: 782–6.

Craven J, Rodin G, Littlefield C (1988). "The Beck Depression Inventory as a screening device for major depression in renal dialysis patients". Int J Psychiatry Med. 18 (4): 365–374.

Cunningham, JL; et al. (2011). "Agreement between physicians' and patients' ratings on the Montgomery Åsberg Depression Rating Scale". J. Affective Disorders. 135 (1-3): 148–53.

De Jonge, P.; Mangano, D.; Whooley, M. A. (2007). "Differential Association of Cognitive and Somatic Depressive Symptoms with Heart Rate Variability in Patients with Stable Coronary Heart Disease: Findings from the Heart and Soul Study". Psychosomatic Medicine. 69 (8): 735–739.

Demyttenaere K, De Fruyt J (2003). "Getting what you ask for: on the selectivity of depression rating scales". Psychotherapy and psychosomatics. 72 (2): 61–70.

Garcia-Esteve L, Ascaso C, Ojuel J, Navarro P (Jun 2003). "Validation of the Edinburgh Postnatal Depression Scale (EPDS) in Spanish mothers". J Affect Disord. 75 (1): 71–6.

Fawzi, M.H., Fawzi, M.M., Abu-Hindi, W., (2012). Arabic version of the Major Depression Inventory as a diagnostic tool: reliability and concurrent and discriminant validity. East. Mediterr. Health J.;18(4):304-10.

Gilbody S, House AO, Sheldon TA (2005). Gilbody, Simon, ed. "Screening and case finding instruments for depression". Cochrane Database of Systematic Reviews (4): CD002792.

Hamilton M (1960). "A Rating Scale for Depression". Journal of Neurology, Neurosurgery and Psychiatry. 23 (1): 56–62.

Hamilton, M (1960) A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry.* 23: 56-62 doi:10.1136/jnnp.23.1.56.

Hamilton M (1966) Assessment of change in psychiatric state by means of rating scales. *Proceedings of the Royal Society of Medicine* 59 (Suppl. 1): 10-13.

Hamilton, M (1967) Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology* 6: 278-96.

Hamilton, M (1969) Standardised assessment and recording of depressive symptoms. *Psychiatria, Neurologia, Neurochirurgia.* 72:201-205.

Hamilton, M (1980) Rating depressive patients. *Journal of Clinical Psychiatry.* 41: 21-24.

Hedlund JL, Viewig BW (1979) The Hamilton rating scale for depression: a comprehensive review. *Journal of Operational Psychiatry* 10:149-165.

Hersen M, Turner SM, Beidel DC (2007). *Adult Psychopathology and Diagnosis (5th ed.).* NY: John Wiley & Sons. *pp. 301–302.*

Katz IR (1998). "Diagnosis and treatment of depression in patients with Alzheimer's disease and other dementias". The Journal of Clinical Psychiatry. 59 Suppl 9: 38–44.

Kriston, L.; Von Wolff, A. (2011). "Not as golden as standards should be: Interpretation of the Hamilton Rating Scale for Depression". Journal of Affective Disorders. 128 (1–2): 175–177.

Kroenke K, Spitzer RL, Williams JB (September 2001). "The PHQ-9: validity of a brief depression severity measure". J Gen Intern Med. 16 (9): 606–13.

Kutcher, Stan (2003). "The Kutcher Adolescent Depression Scale: Assessment of its Evaluative Properties over the Course of an 8-Week Pediatric Pharmacotherapy Trial". Child Adolescent Psychopharmacology. 13 (3): 337–49.

Lesher EL, Berryhill JS. Validation of the Geriatric Depression Scale-Short Form among inpatients. J Clin Psychol. 1994 Mar;50(2):256-60.

Montgomery SA, Asberg M (April 1979). "A new depression scale designed to be sensitive to change". *British Journal of Psychiatry. 134 (4): 382–9.*

Moore MJ, Moore PB, Shaw PJ (October 1998). "Mood disturbances in motor neurone disease". *Journal of the neurological sciences. 160 Suppl 1: S53–6.*

Olsen LR, Jensen DV, Noerholm V, Martiny K, Bech P. The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. Psychol Med 2003; 33, 351-356.

Pedersen, Darlene D. (1 August 2013). *Psych Notes: Clinical Pocket Guide. F.A. Davis. pp. 207ff.*

Raskin A, Schulterbrandt J, Reatig N, McKeon JJ (January 1969). "Replication of factors of psychopathology in interview, ward behavior and self-report ratings of hospitalized depressives". *J. Nerv. Ment. Dis. 148 (1): 87–98.*

Richter, P; J Werner; A Heerlein; A Kraus; H Sauer (1998). "On the validity of the Beck Depression Inventory. A review". *Psychopathology. 31 (3): 160–8.*

Sharp LK, Lipsky MS (September 2002). "Screening for depression across the lifespan: a review of measures for use in primary care settings". *American Family Physician. 66 (6): 1001–8.*

Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. Clin Gerontol. 1986 June;5(1/2):165-173.

Spitzer RL, Kroenke K, Williams JB (November 1999). "Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire". *Journal of the American Medical Association. 282 (18): 1737–44.*

Steer RA, Ball R, Ranieri WF, Beck AT (January 1999). "Dimensions of the Beck Depression Inventory-II in clinically depressed outpatients". *Journal of clinical psychology. 55 (1): 117–28.*

Steer RA, Cavalieri TA, Leonard DM, Beck AT (1999). "Use of the Beck Depression Inventory for Primary Care to screen for major depression disorders". *General hospital psychiatry. 21 (2): 106–11.*

Storch EA, Roberti JW, Roth DA (2004). "Factor structure, concurrent validity, and internal consistency of the Beck Depression Inventory-Second Edition in a sample of college students". *Depression and Anxiety. 19 (3): 187–9.*

Svanborg, P; Åsberg, M (2001). "A comparison between the Beck Depression Inventory (BDI) and the self-rating version of the Montgomery Åsberg Depression Rating Scale (MADRS)". *J. Affective Disorders. 64 (2-3): 203–216.*

Vivilaki VG, Dafermos V, Kogevinas M, Bitsios P, Lionis C (Sep 2009). "The Edinburgh Postnatal Depression Scale: translation and validation for a Greek sample". *BMC Public Health. 9: 329.*

Wechsler H, Grosser GH, Busfield BL Jr (1963) The depression rating scale: a quantitative approach to the assessment of depressive symptomatology. *Archives of General Psychiatry.* 9: 334-343.

Whooley MA, Avins AL, Miranda J, Browner WS (July 1997). "Case-Finding Instruments for Depression: Two Questions Are as Good as Many". *J Gen Intern Med. 12 (7): 439–45.*

Williams JBW (1989) A structured interview guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry* 45: 742-747.

Williams, J. B. W.; Kobak, K. A. (2008). "Development and reliability of a structured interview guide for the Montgomery-Asberg Depression Rating Scale (SIGMA)". *The British Journal of Psychiatry. 192 (1): 52–58.*

Yesavage JA (1988). "Geriatric Depression Scale". *Psychopharmacology Bulletin. 24 (4): 709–11.*

Yesavage JA, Brink TL, Rose TL, et al. (1982-83). Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res.17(1):37-49.

Zigmond, AS; Snaith, RP (1983). "The hospital anxiety and depression scale". *Acta Psychiatrica Scandinavica. 67 (6): 361–370.*

Zimmerman M, Chelminski I, McGlinchey JB, Posternak MA (2008). "A clinically useful depression outcome scale". *Compr Psychiatry. 49 (2): 131–40.*

Zung WW (January 1965). "A self-rating depression scale". *Arch. Gen. Psychiatry. 12: 63–70.*