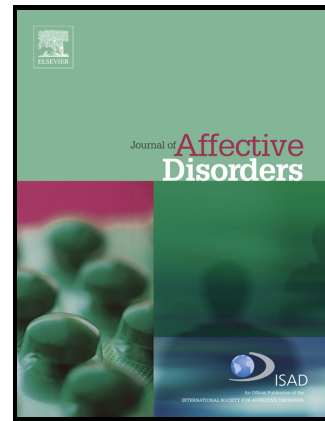


Author's Accepted Manuscript

Depression rating scales in Parkinson's disease: A critical review updating recent literature

Elizabeth Torbey, Nancy A. Pachana, Nadeeka N.W. Dissanayaka



PII: S0165-0327(15)00365-1
DOI: <http://dx.doi.org/10.1016/j.jad.2015.05.059>
Reference: JAD7495

To appear in: *Journal of Affective Disorders*

Received date: 28 October 2014
Revised date: 27 May 2015
Accepted date: 29 May 2015

Cite this article as: Elizabeth Torbey, Nancy A. Pachana and Nadeeka N.W. Dissanayaka, Depression rating scales in Parkinson's disease: A critical review updating recent literature, *Journal of Affective Disorders*, <http://dx.doi.org/10.1016/j.jad.2015.05.059>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Depression rating scales in Parkinson's disease: A critical review updating recent literature

Elizabeth Torbey^{a,b}, BPsychSci, Nancy A. Pachana^a, PhD, Nadeeka N. W. Dissanayaka^{a,b,c*}, PhD

^aSchool of Psychology, The University of Queensland, Brisbane, Australia

^bThe University of Queensland, UQ Centre for Clinical Research, Brisbane, Australia.

^cNeurology Research Centre, Royal Brisbane & Women's Hospital, Brisbane, Australia

*Corresponding Author

Nadeeka N. W. Dissanayaka

Phone: +61 7 33465577, +61 405715622

Email: n.dissanayaka@uq.edu.au

Postal address:

The University of Queensland, UQ Centre for Clinical Research,

Building 71/918 Royal Brisbane & Women's Hospital,

Herston QLD4029, Brisbane, Australia

Accepted manuscript

Abstract

Depression is a prominent non-motor symptom in Parkinson's disease (PD). Assessing depression in PD remains a challenge due to the overlap of somatic symptoms between depression and PD. Other neuropsychiatric manifestations associated with PD, such as cognitive decline, also complicate assessment of depression. Therefore it is critical to investigate the validity of depression rating scales for use in PD. This will allow evaluation of observer- and self- report instruments to be administered in neurologically ill geriatric populations such as PD, and identification of appropriate scales to use in cognitively challenged PD patients. The present review includes all studies examining the validity of depression rating scales in PD. It discusses the usefulness of thirteen depression rating scales in PD. The clinician-rated and widely used HAMD-17 and the self-report GDS scales are recommended for screening and measuring severity of depression in PD. The GDS-15 may be a preferred choice due to its brevity and ease of use design for older adults. Other valid and reliable instruments to use in PD include self-rated scales, such as the HADS-D, HDI, and the BDI, and the observer-report, MADRS. The CSDD displayed satisfactory validity and reliability for identification of PD patients with and without dementia. The PHQ-2, PHQ-10, SDS, CES-D, UPDRS-Depression item, IDS-SR, and IDS-C each showed some evidence of validity or reliability, however further research on the psychometric properties of these scales when used in a PD population are required.

Key words: Depression, Parkinson's Disease, Rating scales, Review

Introduction

With many depression scales available for clinical use in psychiatric populations, it is important to investigate which scales are valid and appropriate for use in Parkinson's disease (PD). It has been well established that several symptoms of depressive disorders overlap with other non-motor symptoms of PD (Gallagher and Schrag, 2012; Wishart and Macphee, 2011). For example, somatic and neurovegetative difficulties such as fatigue, psychoagitation, impaired concentration, and insomnia are seen in both depression and PD. Overlap in symptoms is likely to cause difficulties in the accurate identification and diagnosis of depression in PD, hence contributing to both under-detection of cases as well as under-treatment. The majority of rating scales used to assess depression consist of such overlapping symptoms, and therefore it is important to examine the validity of the use of depression rating scales in PD. For consistency of assessment and to enhance specificity, it has been suggested to use an "inclusive" approach when assessing depression in PD (Marsh et al., 2006), which involves rating of the presence or severity of the symptom regardless of the overlap with PD or other medical conditions.

In clinical practice, patients undergoing assessment for psychiatric disorders such as depression should be interviewed using a standardised clinical interview based on diagnostic criteria, such as those from the Diagnostic Statistical Manual, Fifth Edition (DSM-5) (American Psychiatric Association, 2013). However, rating scales (either self-report, clinician-based, or informant-rated) are often used to screen for psychiatric symptoms and their severity and assist with eventual diagnosis. Often administration of rating scales is also more feasible than conducting interviews for the assessment of psychological disorders in epidemiological studies, surveys, and clinical trials. The use of brief, valid rating scales administered by clinicians and researchers is therefore vital in improving the detection of depressive disorders, which are highly prevalent in PD (Reijnders et al., 2008).

Despite measuring the same overall psychological construct of depression, each rating scale is unique in what symptoms their items aim to assess. The aim of the present review is therefore to explore tools used to measure depression in PD to determine their reliability and validity in PD, in order to ascertain the most useful rating scales. A similar review was previously published by Schrag et al. in 2007. However since then a number of original articles focusing on depression rating scales in PD have been published. For example an original study by Williams et al. 2012 compared the utility of 9 depression rating scales in PD. This review is a comprehensive update of the literature pertaining to reliability and validity studies of all depression rating scales in PD, and provides as well an updated overview, based on this literature, of the utility of these measures in PD.

A literature search was performed using PsycInfo, PubMed and Web of Science databases. The search terms included Parkinson* disease, psychiatric, depress*, assessment, scales, and valid*. No years were specified in the search, and therefore all years up until the present were included. The inclusion criterion for the literature review was review articles and studies that have investigated the validation of depression instruments in PD, and written in English. The results of the search revealed thirteen depression rating scales that have been used in PD. The structure of this review consists two parts for each scale. First is to provide general information about the scale and second is to comprehensively discuss validity and reliability details relating to PD. First part (general information) guides the reader to understand more about the scale when used in the general population, specially those who are unfamiliar with various rating scales used to measure depression. Second part clearly describes studies examining the validity and reliability in PD. For each scale, a brief conclusions as to whether the evidence suggests that the scale is appropriate for use in PD are

outlined. A summary of each depression scale reviewed and its usefulness in PD is provided in Table 4.

Frequently used general depression rating scales

Hamilton Depression Rating Scale (HAMD)

The HAMD was one of the first semi-structured interview measures developed for the clinical evaluation of depression in adults and remains the most widely used measure in clinical practice (Hamilton, 1960). There are multiple versions of the clinician-rated HAMD available, including 6-item, 17-item, 21-item, and 24-item scales (Serrano-Duenas and Soledad Serrano, 2008; Weintraub et al., 2006). The 17-item version (HAMD-17) is the most frequently used version. Each item is scored on a 3-point or 5-point scale, with higher scores indicating greater severity of symptoms. The HAMD exhibits good discriminant validity, test-retest reliability, inter-rater reliability, and good sensitivity to change in non-PD depressed patients (Bagby et al., 2004). One main criticism of the HAMD, however, is that somatic symptoms of depression are heavily represented in item content.

The reliability and validity of the HAMD in PD has been evaluated in a number of studies. The results from studies which have assessed the discriminant validity of the HAMD-17 and HAMD-24 in PD are summarised in Table 1. An optimal HAMD-24 cut-off score for distinguishing between patients with and without a depressive disorder was found to be 9/10, with a high area under the curve (AUC) (0.91) indicating excellent discrimination (Weintraub et al., 2006). In this study, a depressive disorder indicated a diagnosis of major or minor depression according to the gold standard DSM-IV diagnostic criteria. Limitations of the study, however, include the relatively older mean age of patients in the sample (72 years) and the fact that it was a predominantly male sample. The researchers also did not include any cognitive assessment or exclude patients diagnosed with dementia. Leentjens et al. (2000a) investigated the discriminant validity of the HAMD-17, with results also indicating a high AUC (0.95) and suggesting an optimal cut-off score of 13/14. This study did not specify what a diagnosis of 'depressive disorder' included. To assess major depression an optimal cut-off of 12/13 was suggested by two studies (McDonald et al., 2006; Naarding et al., 2002), while Dissanayaka et al. (2007) suggested a higher cut-off of 18/19. The prevalence of major depression was lower in Dissanayaka et al. 2007 study, and this study suggested a cut-off of 12/13 for depressive disorder (major depression, minor depression and dysthymia). A lower optimal cut-off of 6/7 was recently suggested by Williams et al. (2012). The inclusion of depressive subtypes resulting in a very high prevalence of depressive symptomatology (93%) may account for the low optimal cut-off value.

Due to the overlap of somatic symptoms in both depression and PD, Reijnders et al. (2010) investigated the discriminant validity of the HAMD-17 and a modified version, in which items assessing somatic symptoms were excluded. The results showed that the modified version was reduced in specificity. It was therefore recommended that the original HAMD-17 is selected for assisting with diagnosing depression in PD, while an abbreviated version might be more appropriate for screening purposes. Overall, the HAMD-17 has displayed good validity and reliability, and is generally recommended for screening and measuring severity of depression in PD.

Hamilton Depression Inventory (HDI)

The HDI is a self-rated version of the clinician-rated HAMD, and assesses the severity of depressive symptoms over the past two weeks (Dissanayaka et al., 2007; Reynolds and Kobak, 1995). There is a 23-item, 17-item, and 9-item version of this measure, however the 17-item scale remains analogous to the standard 17-item HAMD. Item scores on the HDI-

17 can range from 0 to 2 or 0 to 4. The HDI has been validated and used in the general population, demonstrating high levels of reliability, content validity, criterion validity, construct validity, and clinical efficacy (Reynolds and Kobak, 1995). In PD, an optimal cut-off score of 13.5/14.0 was suggested to discriminate between patients with and without depressive disorder (major depression, minor depression, or dysthymia according to the DSM-IV criteria) and 15.5/16 for major depression (Dissanayaka et al., 2007). The internal consistency of the HDI-17 was also satisfactory (Cronbach's $\alpha = .85$). Overall, the HDI-17 appears to be an appropriate scale to screen and diagnose depressive disorders in PD.

Beck Depression Inventory (BDI)

The BDI is a 21-item self-report assessment tool used to screen, diagnose, and measure severity of depression (Beck et al., 1961; Leentjens et al., 2000b). It is one of the most frequently used self-rated instruments for the assessment of major depression. There exist several adaptations of the BDI in accordance with DSM-IV criteria. The BDI-II, the second revision of the original BDI, included the addition of items to assess agitation, concentration difficulties, and loss of energy (Beck et al., 1996). There has been criticism that the BDI contains several somatic items, however the main focus of BDI items is on psychological symptoms of depression (Schrag et al., 2007). Each of the BDI items is scored from 0 to 3. The BDI has exhibited good test-retest reliability, internal consistency, concurrent validity, and discriminant validity in a variety of patient groups (Beck et al., 1961; Richter et al., 1998).

Studies which have assessed the discriminant validity of the BDI in PD are summarised in Table 2. Leentjens et al. (2000b) examined the validity of the BDI in a sample of 53 PD patients who were diagnosed with or without depression according to the SCID. Results indicated that the optimal cut-off score to discriminate between depressed and non-depressed patients was 13/14 (AUC = 0.86), although it was concluded that the BDI should not be used to dichotomise based on a single cut-off score. A limitation of this study was that they did not specify what was included in the diagnosis of a 'depressive disorder'.

Visser et al. (2006) also reported an adjusted cut-off score of 14/15, with a high AUC (0.88) indicating good discriminant validity of the BDI for a diagnosis of major depression in PD. The test-retest reliability ($r = 0.89$) and internal consistency (Cronbach's $\alpha = 0.88$) were also good. Further studies have demonstrated that the BDI has good reliability and validity for the assessment of major depression or depressive disorders in PD (Levin et al., 1988; Schneider et al., 2010; Silberman et al., 2006; Williams et al., 2012). Overall, the BDI displays good validity and reliability for use in PD, and is recommended for screening and measuring severity of depression in this population.

Hospital Anxiety and Depression Scale – Depression Subscale (HADS-D)

The HADS-D is a 7-item self-report subscale of the HADS which assesses symptoms of depression (Zigmond and Snaith, 1983). It was designed to assess the emotional aspects of depression (e.g. anhedonia) and excludes physical and cognitive symptoms, and suicidal ideation (Schrag et al., 2007). Respondents are required to rate the items on a 4-point scale, ranging from 0 to 3. In the assessment of non-PD patients, the HADS-D has displayed good internal consistency, test-retest reliability, and sensitivity to change in treatment (Bjelland et al., 2002; Herrmann, 1997). The discriminant and concurrent validity of the HADS-D were also satisfactory, however it has been criticised for having poor face validity due to exclusion of the items which assess the most severe symptoms of depression (Bjelland et al., 2002; Schrag et al., 2007).

The discriminant validity of the HADS-D was evaluated against the HAMD-17 in non-demented PD (Mondolo et al., 2006). A cut-off score to best distinguish between

depressed and non-depressed patients (according to HAMD-17 score) was 10/11, with a very high AUC (0.98). HADS-D has displayed good internal consistency and test-retest reliability (Marinus et al., 2002; Rodriguez-Blazquez et al., 2009), and appears to be a reliable instrument for screening depression in PD; however further research to assess the discriminant validity of the scale against DSM-5 criteria is required.

Zung Self-Rating Depression Scale (SDS)

The SDS is a 20-item self-rated scale that screens for and measures the severity of psychological and somatic symptoms of depression (Schrag et al., 2007; Zung, 1965). Items on the scale represent most of the DSM-IV criteria for major depression, however there are a large number of somatic items that may overlap with symptoms from other health conditions such as PD (Schrag et al., 2007). Respondents are required to rate the frequency that each symptom affects their life, with items scored on a 4-point scale ranging from 1 = *a little of the time* to 4 = *most of the time*. There have been multiple shortened versions of the SDS, although the original remains the most frequently used. The SDS has displayed good internal consistency, test-retest reliability, content validity, and criterion validity, and acceptable concurrent validity (Kaneda, 1999; Zung et al., 1965).

The Portuguese Brazilian version of the SDS appears to be valid and reliable to use in non-demented PD patients (Chagas et al., 2010). An optimal cut-off of 54/55 has shown to distinguish between depressed and non-depressed patients against the DSM-IV criteria, with a high AUC (0.93). It has displayed good concurrent validity with the GDS-15 ($\rho = 0.65$), however the internal consistency was moderate (Cronbach's $\alpha = 0.73$). The original English version of the SDS has been used to assess depression in PD, however this version of the scale has not been validated for use in PD (Kanda et al., 2008).

Centre for Epidemiologic Studies Depression Scale (CES-D)

The CES-D is a brief, 20-item self-report scale that was developed to measure depressive symptomatology (Radloff, 1977). The items were derived from other validated depression scales to form a screening instrument for depression, particularly in older adults with physical illness. Respondents are required to indicate the frequency which they have experienced each depressive symptom over the past week on a 4-point scale, ranging from 0 = *rarely or none at all* to 3 = *most or all of the time*. There are multiple versions of the CES-D available, including a short adaptation for older adults, with all versions sharing similar psychometric properties (Andresen et al., 1994). The CES-D is reported to exhibit high internal consistency, adequate test-retest reliability, and good construct and discriminant validity (Parikh et al., 1988; Radloff, 1977). Williams et al. (2012) is the only study to investigate the validity and reliability of the CES-D in PD patients without cognitive decline. The results indicated that the internal consistency was high (Cronbach's $\alpha = 0.92$) and an optimal cut-off score was 11/12, with an AUC of 0.79. Overall, further research is required to investigate the reliability and validity of the CES-D for screening depression in PD.

Montgomery-Asberg Depression Rating Scale (MADRS)

The MADRS is a 10-item clinician-rated scale and represent the full DSM-IV criteria of a major depressive episode, with the exception of hypersomnia, increased appetite, and psychomotor retardation or agitation (Montgomery and Asberg, 1979). It has exhibited good face validity, criterion validity, and concurrent validity, and high inter-rater reliability and internal consistency (Davidson et al., 1986; Montgomery and Asberg, 1979). While Leentjens et al. (2000a) found an optimal cut-off score of 14/15, with a high AUC (0.90), Silberman et al. (2006) showed a lower cut-off of 9/10 with a satisfactory AUC (0.84). Both studies did not specify what constituted a diagnosis of 'depression'. The results of Williams et al. (2012)

indicated an AUC of 0.88 and adequate internal consistency (Cronbach's $\alpha = 0.83$). As with the HAMD, Reijnders et al. (2010) investigated the discriminant validity of the original MADRS as well as a modified version with somatic symptoms excluded. It was concluded that the original version of the MADRS should be selected for assisting with diagnosing depression in PD and the abbreviated version of the MADRS might be more appropriate for brief screening. Overall, the MADRS appears to be a valid clinician-rated scale for the assessment of depression in PD.

Infrequently used general depression rating scales

Patient Health Questionnaires (PHQ-2 and PHQ-9)

The PHQ-9 is a self-rated assessment tool and is based on the DSM-IV criteria for major depression (Chagas et al., 2013; Kroenke et al., 2001). The items are scored on a 4-point Likert scale, ranging from 0 = *not at all* to 3 = *nearly every day*. The PHQ-2 is a brief, clinician-rated scale and the two items assess whether patients have experienced 'loss of interest or pleasure' and 'feeling down or hopeless' (Chagas et al., 2011). In the general population, the PHQ-9 has exhibited good internal consistency, test-retest reliability, and criterion (discriminant) validity, while the PHQ-2 has also displayed satisfactory construct and criterion validity (Kroenke et al., 2001, 2003). In PD, both PHQ-9 and PHQ-2 have demonstrated high discriminant validity, and the optimal cut-off values to discriminate major depression were 8/9 and 2/3, respectively (Chagas et al., 2011; Chagas et al., 2013). An optimal cut-off of 5/6 for PHQ-9 has shown to appropriately discriminate between all subtypes of depression (Williams et al., 2012). The PHQ-9 has depicted adequate internal consistency, and moderate concurrent validity against the GDS-15 and SDS (Chagas et al., 2013; Williams et al., 2012).

Inventory of Depression Symptoms: Clinical and Self-Report (IDS-C and IDS-SR)

IDS scales assess for symptoms of major depression, and focus on frequency of symptoms, as opposed to intensity of symptoms. There are 28-item and, more recently, 30-item versions of both of the IDS scales. Both item versions of the IDS-C and IDS-SR have demonstrated good psychometric properties, including adequate face validity, internal consistency, inter-rater reliability, concurrent validity, and discriminant validity (Rush et al., 1986; Rush et al., 1996). With better item coverage, the 30-item version is suggested for the evaluation of depressive symptom severity. In PD, IDS-C and IDS-SR have demonstrated good internal consistency, and optimal cut-off values to discriminate between patients with and without major depression were 11/12 and 13/14, respectively (Williams et al., 2012).

World Health Organisation (WHO)-Five Wellbeing Index (WHO-5)

The WHO-5 is a brief tool developed to measure the level of emotional well-being over a 2 week period (Schneider et al., 2010). Items focus on positive affect, and therefore lower scores indicate greater depressive symptoms. The WHO-5 has displayed good psychometric properties as a reliable screening tool for depression (Henkel et al., 2004; Lowe et al., 2004). This scale was suggested as an alternative to the commonly used BDI to quickly screen for low emotional well-being and depression in PD patients. WHO-5 has demonstrated good internal consistency, and an optimal cut-off of 12/13 has been suggested to distinguish PD patients with DSM-IV depressive disorders (Schneider et al., 2010).

Depression rating scales developed for use in older adults

Geriatric Depression Scale (GDS)

The GDS is a brief, self-report measure, which focuses on non-somatic symptoms of depression, such as the psychological aspects (e.g. hopelessness) and social consequences of depression. It avoids inclusion of overlapping somatic symptoms (e.g. insomnia) (Dissanayaka et al., 2011; McDonald et al., 2006; Yesavage et al., 1982). The two commonly used versions of the instrument have 15 or 30-items. All items have a yes or no response set, scored as 0 or 1, and the instrument is easy to use. Psychometric properties of the GDS in depressed older adults without PD have demonstrated good discriminant validity, internal consistency and test-retest reliability. It was highly correlated with other depression rating scales (Schrag et al., 2007; Yesavage et al., 1982).

The studies which have assessed the discriminant validity of the GDS-15 or GDS-30 in PD are summarised in Table 3. An optimal cut-off of 13/14 in GDS-30 was suggested to distinguish PD patients diagnosed with major and minor depression, and non-depressed patients (Ertan et al., 2005), while 9/10 was shown to distinguish patients with and without major depression (McDonald et al., 2006). The GDS-30 has also displayed strong internal consistency (Cronbach's $\alpha = 0.92$) and split-half correlation coefficient ($\rho = 0.91$) (Ertan et al., 2005). Further studies have provided support for the use of both versions of the GDS as a valid screen for depressive symptoms and to diagnose depression in PD (Dissanayaka et al., 2007; Mondolo et al., 2006; Weintraub et al., 2006; Williams et al., 2012). The GDS-15 has been suggested a suitable tool to screen for major or minor depression in PD across all ages (Weintraub et al., 2007).

Depression rating scales developed for use in dementia

Cornell Scale for Depression in Dementia (CSDD)

The CSDD is a clinician-rated instrument developed specifically for the assessment of depression in patients with dementia (Alexopoulos et al., 1988a, b). Scoring is based on observation and interviews with the patient and an informant (frequently the patient's caregiver). Informants are used to provide collateral information. The scale was developed to assess severity of depression, although it is also used for the screening of depression in dementia. Psychometric properties of the CSDD are acceptable, with satisfactory internal consistency, inter-rater reliability, concurrent validity, and sensitivity to change in the context of treatment (Alexopoulos et al., 1988a, b). It has been validated for the use in dementia and cognitively intact geriatric populations (Muller-Thomsen et al., 2005).

There are high prevalence rates of dementia in PD, therefore the CSDD is considered an appropriate and potentially useful tool to assess depression in patients with comorbid cognitive impairment (Schrag et al., 2007; Williams and Marsh, 2009). The content of the scale is also limited in terms of questions that might be associated with motor symptoms of PD, however there is some overlap with PD difficulties (e.g. psychomotor retardation). Williams and Marsh (2009) assessed the psychometric properties of the CSDD in 134 PD patients with and without dementia. Patients were categorised as having a depressive disorder if they met DSM-IV criteria for major depressive disorder, major depressive disorder in partial remission, minor depression, dysthymia, or depressive disorder not otherwise specified. The discriminant validity of the scale to distinguish which patients were identified as having a depressive disorder was moderate, with an optimal cut-off score of 7/8 and AUC of 0.82. Internal consistency of the CSDD was also acceptable (Cronbach's $\alpha = 0.84$). These results indicate that the CSDD is a valid tool for identifying depressive disorders across the cognitive spectrum in PD.

Screening item only for assessment of depression

Unified Parkinson's Disease Rating Scale (UPDRS) Depression item

The UPDRS widely used to assess the severity of PD includes a single observer-report item to assess depressed mood on a 5-point scale ranging from 0 = *normal* to 4 = *severe* (Fahn et al., 1987; Goetz et al., 2008). This item has shown moderate concurrent validity against the HAMD, HADS-D, and GDS (Gallagher et al., 2012; Holroyd et al., 2008). Except for the study by Chagas et al. (2011), other studies have indicated a low discriminant validity for the UPDRS-Depression item (Holroyd et al., 2008; Starkstein and Merello, 2007; Williams et al., 2012). Overall, the UPDRS-Depression item does not currently appear to be valid as a screening tool for depression.

Author declaration

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. Dr Nadeeka Dissanayaka and Prof Nancy Pachana are employed by the University of Queensland. Dr Dissanayaka's salary is supported by the Lions Medical Research Foundation Fellowship, and this is stated in the acknowledgments.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). She is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author and which has been configured to accept email from n.dissanayaka@uq.edu.au.

Conclusion

This review has provided an updated summary regarding the validity of existing depression rating scales for use with patients with PD. The results will benefit clinicians who require assessment of affective disorders in patients with PD, as it can assist them in choosing an appropriate scale for screening or to assist with the diagnosis of psychiatric disorders in this population. In situations where there are several scales with similar psychometric properties which are appropriate, it is understandable that a clinician or researcher would consider factors such as cost and convenience in selection of such a screening scale (Schrag and Leentjens, 2012). A summary of all depression rating scales discussed in this review is displayed in Table 4.

Of the depression rating scales, the clinician-rated HAMD-17 displayed very good reliability and discriminant validity, and is recommended for screening and measuring severity of depression in PD. The optimal cut-off to discriminate patients with and without

depressive disorder for the HAMD-17 was 12/13 across a number of studies (Dissanayaka et al., 2007; McDonald et al., 2006; Naarding et al., 2002). The clinician-rated MADRS also exhibited very good validity and reliability in PD, although it did not appear to perform as well as the HAMD-17 in identifying depression in PD patients (Leentjens et al., 2000a). Reijnders et al. (2010) concluded, however, that the HAMD-17 and MADRS are suitable for diagnosing depression in PD, but that abbreviated versions of these scales, which exclude somatic symptoms, might be more appropriate for screening purposes.

The self-rated GDS-15 and GDS-30 were both found to have good validity and reliability and are recommended as screening tools for depression in PD. While both scales are designed for easy use in older adults who might have sensory or other impairments, clinicians may prefer to use the GDS-15 given it is shorter in length. The GDS-15 also appears to be appropriate for use in PD patients of all ages (Weintraub et al., 2007).

Other self-rated scales which were found to display good validity and reliability for the use of screening or diagnosing severity of depression in PD included the HADS-D, HDI, and the BDI. The HADS-D also displayed sound test-retest reliability. The clinician-rated PHQ-2 and self-rated PHQ-9 were both found to be brief, valid, and reliable tools for screening depression in PD, as was the WHO-5. The Portuguese Brazilian version of the SDS also exhibited good validity and reliability as a screen for depression in PD, however more research on the original English version of the scale is required. For PD patients who appear to have cognitive impairment, the CSDD is recommended for use as it displayed satisfactory validity and reliability for identifying depression in patients with and without dementia. The CES-D, UPDRS-Depression item, IDS-SR, and IDS-C each showed some evidence of validity or reliability, however this was primarily investigated in the study by Williams et al. (2012), the results of which are likely to be impacted by recruitment bias. Further research on the psychometric properties of each of these scales is therefore required. We note that none of the original articles reviewed reported responsiveness, and an investigation of such information can also be a focus for future studies.

Contributors

Elizabeth Torbey: Conception, Organisation, Writing of the first draft, Review and critique

Nancy Pachana: Conception, Review and critique

Nadeeka Dissanayaka: Conception, Organisation, Writing of the first draft, Review and critique

All authors have approved the final article.

Acknowledgment

We thank the Lions Medical Research Foundation for supporting Dr Nadeeka Dissanayaka's fellowship.

Role of the funding source

Lions Medical Research Foundation supported Dr Dissanayaka's salary.

References

- Alexopoulos, G.S., Abrams, R.C., Young, R.C., Shamoian, C.A., 1988a. Cornell Scale for Depression in Dementia. *Biol Psychiat*. 23, 271-284.
- Alexopoulos, G.S., Abrams, R.C., Young, R.C., Shamoian, C.A., 1988b. Use of the Cornell scale in nondemented patients. *J Am Geriatr Soc*. 36, 230-236.
- American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders: DSM-5, 5th ed. American Psychiatric Association, Washington, DC.
- Andresen, E.M., Malmgren, J.A., Carter, W.B., Patrick, D.L., 1994. Screening for depression in well older adults: Evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *Am J Prev Med*. 10, 77-84.
- Bagby, R.M., Ryder, A.G., Schuller, D.R., Marshall, M.B., 2004. The Hamilton Depression Rating Scale: Has the gold standard become a lead weight? *Am J Psychiat*. 161, 2163-2177.
- Beck, A.T., Steer, R.A., Brown, G.K., 1996. Manual for the Beck Depression Inventory–II. Psychological Corporation, San Antonio, TX.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. *Arch Gen Psychiat*. 4, 561-571.
- Bjelland, I., Dahl, A.A., Haug, T.T., Neckelmann, D., 2002. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res*. 52, 69-77.
- Chagas, M.H., Crippa, J.A., Loureiro, S.R., Hallak, J.E., Meneses-Gaya, C., Machado-de-Sousa, J.P., Rodrigues, G.R., Filho, A.S., Sanches, R.F., Tumas, V., 2011. Validity of the PHQ-2 for the screening of major depression in Parkinson's disease: Two questions and one important answer. *Aging Ment Health*. 15, 838-843.
- Chagas, M.H., Tumas, V., Loureiro, S.R., Hallak, J.E., Trzesniak, C., de Sousa, J.P., Rodrigues, G.G., Santos Filho, A., Crippa, J.A., 2010. Validity of a Brazilian version of the Zung self-rating depression scale for screening of depression in patients with Parkinson's disease. *Parkinsonism Relat D*. 16, 42-45.
- Chagas, M.H., Tumas, V., Rodrigues, G.R., Machado-de-Sousa, J.P., Filho, A.S., Hallak, J.E., Crippa, J.A., 2013. Validation and internal consistency of Patient Health Questionnaire-9 for major depression in Parkinson's disease. *Age Ageing*. 42, 645-649.
- Davidson, J., Turnbull, C.D., Strickland, R., Miller, R., Graves, K., 1986. The Montgomery-Asberg Depression Scale: Reliability and validity. *Acta Psychiat Scand*. 73, 544-548.
- Dissanayaka, N., O'Sullivan, J.D., Silburn, P.A., Mellick, G.D., 2011. Assessment methods and factors associated with depression in Parkinson's disease. *J Neurol Sci*. 310, 208-210.
- Dissanayaka, N., Sellbach, A., Matheson, S., Marsh, R., Silburn, P.A., O'Sullivan, J.D., Byrne, G.J., Mellick, G.D., 2007. Validity of Hamilton Depression Inventory in Parkinson's disease. *Movement Disord*. 22, 399-403.
- Ertan, F.S., Ertan, T., Kiziltan, G., Uygucgil, H., 2005. Reliability and validity of the Geriatric Depression Scale in depression in Parkinson's disease. *J Neurol Neurosur Ps*. 76, 1445-1447.
- Fahn, S., Elton, R.L., UPDRS Program Members, 1987. Unified Parkinson's disease rating scale, In: Fahn, S., Marsden, C.D., Goldstein, M., Calne, D.B. (Eds.), *Recent developments in Parkinson's disease*. Macmillan Healthcare Information, Florham Park, pp. 153–163.
- Gallagher, D.A., Goetz, C.G., Stebbins, G., Lees, A.J., Schrag, A., 2012. Validation of the MDS-UPDRS Part I for nonmotor symptoms in Parkinson's disease. *Movement Disord*. 27, 79-83.
- Gallagher, D.A., Schrag, A., 2012. Psychosis, apathy, depression and anxiety in Parkinson's disease. *Neurobiol Dis*. 46, 581-589.
- Goetz, C.G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M.B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J.,

- Kulisevsky, J., Lang, A.E., Lees, A., Leurgans, S., LeWitt, P.A., Nyenhuis, D., Olanow, C.W., Rascol, O., Schrag, A., Teresi, J.A., Hilten, J.J., LaPelle, N., 2008. Movement Disorder Society-Sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disord.* 23, 2129-2170.
- Hamilton, M., 1960. A rating scale for depression. *J Neurol Neurosurg Ps.* 23, 56-62.
- Henkel, V., Mergl, R., Coyne, J.C., Kohlen, R., Moller, H.J., Hegerl, U., 2004. Screening for depression in primary care: Will one or two items suffice? *Eur Arch Psy Clin N.* 254, 215-223.
- Herrmann, C., 1997. International experiences with the Hospital Anxiety and Depression Scale--a review of validation data and clinical results. *J Psychosom Res.* 42, 17-41.
- Holroyd, S., Currie, L.J., Wooten, G.F., 2008. Validity, sensitivity and specificity of the mentation, behavior and mood subscale of the UPDRS. *Neurol Res.* 30, 493-496.
- Kanda, F., Oishi, K., Sekiguchi, K., Kuga, A., Kobesho, H., Shirafuji, T., Higuchi, M., Ishihara, H., 2008. Characteristics of depression in Parkinson's disease: Evaluating with Zung's Self-Rating Depression Scale. *Parkinsonism Relat D.* 14, 19-23.
- Kaneda, Y., 1999. Usefulness of the Zung self-rating depression scale for schizophrenics. *J Med Invest.* 46, 75-78.
- Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med.* 16, 606-613.
- Kroenke, K., Spitzer, R.L., Williams, J.B., 2003. The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Med Care.* 41, 1284-1292.
- Leentjens, A.F., Verhey, F.R., Lousberg, R., Spitsbergen, H., Wilming, F.W., 2000a. The validity of the Hamilton and Montgomery-Asberg depression rating scales as screening and diagnostic tools for depression in Parkinson's disease. *Int J Geriatr Psych.* 15, 644-649.
- Leentjens, A.F., Verhey, F.R., Luijckx, G.J., Troost, J., 2000b. The validity of the Beck Depression Inventory as a screening and diagnostic instrument for depression in patients with Parkinson's disease. *Movement Disord.* 15, 1221-1224.
- Levin, B.E., Llabre, M.M., Weiner, W.J., 1988. Parkinson's disease and depression: Psychometric properties of the Beck Depression Inventory. *J Neurol Neurosurg Ps.* 51, 1401-1404.
- Lowe, B., Spitzer, R.L., Grafe, K., Kroenke, K., Quenter, A., Zipfel, S., Buchholz, C., Witte, S., Herzog, W., 2004. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *J Affect Disorders.* 78, 131-140.
- Marinus, J., Leentjens, A.F., Visser, M., Stiggelbout, A.M., van Hilten, J.J., 2002. Evaluation of the hospital anxiety and depression scale in patients with Parkinson's disease. *Clin Neuropharmacology.* 25, 318-324.
- Marsh, L., McDonald, W.M., Cummings, J., Ravina, B., 2006. Provisional diagnostic criteria for depression in Parkinson's disease: report of an NINDS/NIMH Work Group. *Movement Disord.* 21, 148-158.
- McDonald, W.M., Holtzheimer, P.E., Haber, M., Vitek, J.L., McWhorter, K., DeLong, M., 2006. Validity of the 30-item geriatric depression scale in patients with Parkinson's disease. *Movement Disord.* 21, 1618-1622.
- Mondolo, F., Jahanshahi, M., Grana, A., Biasutti, E., Cacciatori, E., Di Benedetto, P., 2006. The validity of the hospital anxiety and depression scale and the geriatric depression scale in Parkinson's disease. *Behav Neurol.* 17, 109-115.
- Montgomery, S.A., Asberg, M., 1979. A new depression scale designed to be sensitive to change. *Brit J Psychiat.* 134, 382-389.
- Muller-Thomsen, T., Arlt, S., Mann, U., Mass, R., Ganzer, S., 2005. Detecting depression in Alzheimer's disease: Evaluation of four different scales. *Arch Clin Neuropsych.* 20, 271-276.

- Naarding, P., Leentjens, A.F., van Kooten, F., Verhey, F.R., 2002. Disease-specific properties of the Rating Scale for Depression in patients with stroke, Alzheimer's dementia, and Parkinson's disease. *J Neuropsych Clin N.* 14, 329-334.
- Parikh, R.M., Eden, D.T., Price, T.R., Robinson, R.G., 1988. The sensitivity and specificity of the Center for Epidemiologic Studies Depression Scale in screening for post-stroke depression. *Int J Psychiat Med.* 18, 169-181.
- Radloff, L.S., 1977. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psych Meas.* 1, 385-401.
- Reijnders, J.S., Ehrt, U., Weber, W.E., Aarsland, D., Leentjens, A.F., 2008. A systematic review of prevalence studies of depression in Parkinson's disease. *Mov Disord.* 23, 183-189.
- Reijnders, J.S., Lousberg, R., Leentjens, A.F., 2010. Assessment of depression in Parkinson's disease: The contribution of somatic symptoms to the clinimetric performance of the Hamilton and Montgomery-Asberg rating scales. *J Psychosom Res.* 68, 561-565.
- Reynolds, W.M., Kobak, K.A., 1995. Reliability and validity of the Hamilton Depression Inventory: A paper-and-pencil version of the Hamilton Depression Rating Scale Clinical Interview. *Psychol Assessment.* 7, 472-483.
- Richter, P., Werner, J., Heerlein, A., Kraus, A., Sauer, H., 1998. On the validity of the Beck Depression Inventory. A review. *Psychopathology.* 31, 160-168.
- Rodriguez-Blazquez, C., Frades-Payo, B., Forjaz, M.J., de Pedro-Cuesta, J., Martinez-Martin, P., 2009. Psychometric attributes of the Hospital Anxiety and Depression Scale in Parkinson's disease. *Movement Disord.* 24, 519-525.
- Rush, A.J., Giles, D.E., Schlessner, M.A., Fulton, C.L., Weissenburger, J., Burns, C., 1986. The Inventory for Depressive Symptomatology (IDS): Preliminary findings. *Psychiat Res.* 18, 65-87.
- Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B., Trivedi, M.H., 1996. The Inventory of Depressive Symptomatology (IDS): Psychometric properties. *Psychol Med.* 26, 477-486.
- Schneider, C.B., Pilhatsch, M., Rifati, M., Jost, W.H., Wodarz, F., Ebersbach, G., Djundja, D., Fuchs, G., Gies, A., Odin, P., Reifschneider, G., Wolz, M., Bottesi, A., Bauer, M., Reichmann, H., Storch, A., 2010. Utility of the WHO-Five Well-being Index as a screening tool for depression in Parkinson's disease. *Movement Disord.* 25, 777-783.
- Schrag, A., Barone, P., Brown, R.G., Leentjens, A.F., McDonald, W.M., Starkstein, S., Weintraub, D., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G.T., Goetz, C.G., 2007. Depression rating scales in Parkinson's disease: Critique and recommendations. *Movement Disord.* 22, 1077-1092.
- Schrag, A., Leentjens, A.F., 2012. Parkinson disease: Scales to detect depression in Parkinson disease. *Nat Rev Neurol.* 8, 359-360.
- Serrano-Duenas, M., Soledad Serrano, M., 2008. Concurrent validation of the 21-item and 6-item Hamilton Depression Rating Scale versus the DSM-IV diagnostic criteria to assess depression in patients with Parkinson's disease: An exploratory analysis. *Parkinsonism Relat Disord.* 14, 233-238.
- Silberman, C.D., Laks, J., Capitaio, C.F., Rodrigues, C.S., Moreira, I., Engelhardt, E., 2006. Recognizing depression in patients with Parkinson's disease: Accuracy and specificity of two depression rating scale. *Arq Neuro-Psiquiat.* 64, 407-411.
- Starkstein, S.E., Merello, M., 2007. The Unified Parkinson's Disease Rating Scale: Validation study of the mentation, behavior, and mood section. *Movement Disord.* 22, 2156-2161.
- Visser, M., Leentjens, A.F., Marinus, J., Stiggelbout, A.M., van Hilten, J.J., 2006. Reliability and validity of the Beck depression inventory in patients with Parkinson's disease. *Movement Disord.* 21, 668-672.

- Weintraub, D., Oehlberg, K.A., Katz, I.R., Stern, M.B., 2006. Test characteristics of the 15-item geriatric depression scale and Hamilton depression rating scale in Parkinson disease. *Am J Geriatr Psychiat*. 14, 169-175.
- Weintraub, D., Saboe, K., Stern, M.B., 2007. Effect of age on geriatric depression scale performance in Parkinson's disease. *Movement Disord*. 22, 1331-1335.
- Williams, J.R., Hirsch, E.S., Anderson, K., Bush, A.L., Goldstein, S.R., Grill, S., Lehmann, S., Little, J.T., Margolis, R.L., Palanci, J., Pontone, G., Weiss, H., Rabins, P., Marsh, L., 2012. A comparison of nine scales to detect depression in Parkinson disease: Which scale to use? *Neurology*. 78, 998-1006.
- Williams, J.R., Marsh, L., 2009. Validity of the Cornell scale for depression in dementia in Parkinson's disease with and without cognitive impairment. *Movement Disord*. 24, 433-437.
- Wishart, S., Macphee, G.J., 2011. Evaluation and management of the non-motor features of Parkinson's disease. *Ther Adv Chronic Dis*. 2, 69-85.
- Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., Leirer, V.O., 1982. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiat Res*. 17, 37-49.
- Zigmond, A.S., Snaith, R.P., 1983. The hospital anxiety and depression scale. *Acta Psychiat Scand*. 67, 361-370.
- Zung, W.W., 1965. A Self-Rating Depression Scale. *Arch Gen Psychiat*. 12, 63-70.
- Zung, W.W., Richards, C.B., Short, M.J., 1965. Self-rating depression scale in an outpatient clinic. Further validation of the SDS. *Arch Gen Psychiat*. 13, 508-515.

Table 1: A summary of the studies which have assessed the discriminant validity of the HAMD-17 and HAMD-24

Reference	Version	Participants	Diagnosis	AUC	Cut-off scores
Weintraub et al. (2006)	HAMD-24	$N = 148$, mean age 72 years, predominantly male	SCID	0.91	Optimal: 9/10 ^a Screening: 9/10 ^a Diagnostic: 15/16 ^a
Leentjens et al. (2000a)	HAMD-17	$N = 63$, non-demented	SCAN	0.95	Optimal: 13/14 Screening: 11/12 Diagnostic: 16/17
Naarding et al. (2002)	HAMD-17	$N = 85$, non-demented	SCID-D	0.94	Optimal: 12/13 ^c Screening: 9/10 ^c Diagnostic: 15/16 ^c
McDonald et al. (2006)	HAMD-17	$N = 50$, non-demented	SCID	0.82	Optimal: 12/13 ^c
Dissanayaka et al. (2007)	HAMD-17	$N = 79$, non-demented	MINI-Plus	0.96	Optimal: 12/13 ^b Screening: 9/10 ^b Diagnostic: 14/15 ^b Optimal: 18/19 ^c
Williams et al. (2012)	HAMD-17	$N = 229$, non-demented	SCID	0.86	Optimal: 6/7 ^{d,e}

Note. ^aCut-off to differentiate between a diagnosis of depressive disorder (major depression or minor depression) versus no depressive disorder. ^bCut-off to differentiate between a diagnosis of a depressive disorder (major depression, minor depression, or dysthymia) versus no depressive disorder. ^cCut-off to differentiate between diagnosis of major depressive disorder or no major depressive disorder. ^dCut-off to differentiate between diagnosis of depression (major depression, major depressive episode in partial remission, minor depression, dysthymia, adjustment disorder with depressed mood, or depressive disorder not otherwise specified) and absence of depression. ^eWilliams et al. (2012) did not recommend a particular cut-off score for use in clinical practice.

Table 2: A summary of the studies which have assessed the discriminant validity of the BDI, BDI-1A, and BDI-II

Reference	Version	Participants	Diagnosis	AUC	Cut-off scores
Leentjens et al. (2000b)	BDI	$N = 53$, non-demented	SCID-D	0.86	Optimal: 13/14 Screening: 8/9 Diagnostic: 16/17
Visser et al. (2006)	BDI	$N = 92$, non-demented	SCID-D	0.88	Optimal: 14/15 ^a
Silberman et al. (2006)	BDI	$N = 46$, mild-moderate PD severity, non-demented	DSM-IV criteria	0.80	Diagnostic: 17/18
Schneider et al. (2010)	BDI-1A	$N = 209$, non-demented	MINI	0.92 ^b 0.90 ^c	Optimal: 14/15 ^b
Williams et al. (2012)	BDI-II	$N = 229$, non-demented	SCID	0.85	Optimal: 6/7 ^{d,e}

Note. ^aCut-off to differentiate between a diagnosis of major depressive disorder versus no major depressive disorder. ^bCut-off to differentiate between a diagnosis of a depressive disorder (major depression or minor depression) and no depressive disorder. ^cCut-off to differentiate between a diagnosis of a depressive disorder (major depression, minor depression, or dysthymia) and no depressive disorder. ^dCut-off to differentiate between a diagnosis of depression (major depression, major depressive episode in partial remission, minor depression, dysthymia, adjustment disorder with depressed mood, or depressive disorder not otherwise specified) and absence of depression. ^eWilliams et al. (2012) did not recommend specific cut-off scores for use in clinical practice.

Table 3: A summary of the studies which have assessed the discriminant validity of the GDS-15 and GDS-30

Reference	Version	Participants	Diagnosis	AUC	Cut-off scores
Ertan et al. (2005)	GDS-30	$N = 109$, non-demented	DSM-IV criteria	0.89	Optimal: 13/14 ^a Screening: 8/9 or 9/10 ^a Diagnostic: 14/15 or 15/16 ^a
McDonald et al. (2006)	GDS-30	$N = 50$, non-demented	SCID	0.88	Optimal: 9/10 ^b
Mondolo et al. (2006)	GDS-30	$N = 46$, non-demented	HAMD- 17	0.90	Optimal: 10/11 Screening: 10/11 Diagnostic: 12/13
Williams et al. (2012)	GDS-30	$N = 229$, non-demented	SCID	0.83	Optimal: 9/10 ^{d,e}
Weintraub et al. (2006)	GDS-15	$N = 148$, mean age 72 years, predominantly male	SCID	0.92	Optimal: 4/5 ^a Screening: 4/5 ^a Diagnostic: 6/7 ^a
Weintraub et al. (2007)	GDS-15	$N = 58$ (< 65 years) $N = 88$ (65-75 years) $N = 81$ (> 75 years)	SCID	0.92 0.91 .095	Optimal: 4/5 ^a Optimal: 4/5 ^a Optimal: 5/6 ^a
Dissanayaka et al. (2007)	GDS-15	$N = 79$, non-demented	MINI- Plus	0.91	Optimal: 6/7 ^c Screening: 4/5 ^c Diagnostic: 9/10 ^c Optimal: 8/9 ^b

Note.^aCut-off to differentiate between a diagnosis of depressive disorder (major depression or minor depression) versus no depressive disorder. ^bCut-off to differentiate between a diagnosis of major depressive disorder versus no major depressive disorder. ^cCut-off to differentiate between a diagnosis of depressive disorder (major depression, minor depression, or dysthymia) and no depressive disorder. ^dCut-off to differentiate between diagnosis of depression (major depression, major depressive episode in partial remission, minor depression, dysthymia, adjustment disorder with depressed mood, or depressive disorder not otherwise specified) and absence of depression. ^eWilliams et al. (2012) did not recommend specific cut-off scores for use in clinical practice.

Table 4

A Summary of all depression scales reviewed and their usefulness in PD

Scale	Number of validity studies	Usefulness in PD	Key References
Hamilton Depression Scales - 17 & 24 item versions (HAMD-17 and HAMD-24)	7	Good validity and reliability across versions. Recommended for screening and measuring severity of depression in PD.	Leentjens et al. (2000a); Weintraub et al. (2006); Dissanayaka et al. (2007)
Hamilton Depression Inventory (HDI)	1	Good validity and reliability. Recommended for screening and diagnosis of depression in PD.	Dissanayaka et al. (2007)
Beck Depression Inventory (BDI)	6	Good validity and reliability. Recommended for screening and measuring severity of depression in PD.	Leentjens et al. (2000b); Visser et al. (2006); Schneider et al. (2010)
Hospital and Anxiety Depression Scale – Depression subscale (HADS-D)	3	Good reliability, but limitations associated with discriminant validity. Further research is warranted.	Mondolo et al. (2006); Marinus et al. (2002)
Zung Self-rating Depression Scale (SDS) –Brazilian version	1	Brazilian version displayed good reliability and validity for screening of depression in PD. Further research of English version is warranted.	Chagas et al. (2010);
Centre for Epidemiologic Studies Depression Scale (CES-D)	1	Some evidence of reliability and validity, however further research is warranted.	Williams et al. (2012)
Montgomery-Asberg Depression Rating Scale (MADRS)	4	Good discriminant validity. Recommended for assessment of diagnosis of depression in PD.	Leentjens et al. (2000a); Silberman et al. (2006)
Patient Health Questionnaires – 2 & 9 item versions (PHQ-2 and PHQ-9) – Brazilian versions	3	Good discriminant validity across Brazilian versions and recommended as screening tool for depression in PD. Further research of English versions is warranted.	Chagas et al. (2011); Chagas et al. (2013); Williams et al. (2012)

Inventory of Depression Symptoms – Self-Report (IDS-SR) and Clinician (IDS-C)	1	Some evidence of reliability and validity across versions, however further research is warranted.	Williams et al. (2012)
WHO-Five (WHO-5)	1	Good reliability and discriminant validity. Recommended as a brief screening tool for depression in PD.	Schneider et al. (2010)
Geriatric Depression Scales – 15 & 30 item versions (GDS-15 and GDS-30)	7	Good reliability and discriminant validity across versions. Recommended as a screening tool for depression in older PD patients, and possibly PD patients of all ages.	Ertan et al. (2005); McDonald et al. (2006); Dissanayaka et al. (2007); Weintraub et al. (2007)
Cornell Scale for Depression in Dementia (CSDD)	1	Good reliability and discriminant validity. Recommended for identifying depression across cognitive spectrum in PD, but particularly for those with suspected cognitive decline.	Williams and Marsh (2009)
Unified Parkinson's disease Rating Scale (UPDRS)-Depression item	5	Evidence of concurrent validity, however further research is warranted. Mixed findings regarding discriminant validity of original UPDRS depression item.	Gallagher et al. (2012); Starkstein and Merello (2007); Williams et al. (2012)

Highlights

- The present review updates recent literature focussed on depression rating scales in Parkinson's disease (PD)
- It discusses the usefulness of thirteen depression rating scales in PD.
- The clinician-rated and widely used HAMD-17 and the self-report GDS scales are recommended for screening and measuring severity of depression in PD.
- Other valid and reliable instruments to use in PD include self-rated scales, such as the HADS-D, HDI, and the BDI, and the observer-report, MADRS. The CSDD displayed satisfactory validity and reliability for identification of PD patients with and without dementia.
- The PHQ-2, PHQ-10, SDS, CES-D, UPDRS-Depression item, IDS-SR, and IDS-C each showed some evidence of validity or reliability, however further research on the psychometric properties of these scales when used in a PD population are required.