

# Chapter 2

## Rating Scales for Depression

Cristina Cusin, Huaiyu Yang, Albert Yeung,  
and Maurizio Fava

**Abstract** Over the past few decades, a number of clinician-rated and patient-rated instruments have been developed as primary efficacy measures in depression clinical trials. All those scales have relative strengths and weaknesses and some of them have been more successful than others, and have become the gold standards for depression clinical research. With all these measures available and with the evidence of their variable performance in clinical trials, it is becoming increasingly important to select primary efficacy measures that are reliable, valid, and that fit well within the aims of depression clinical trials. This article will review the main considerations that investigators need to make when choosing a primary efficacy measure for major depressive disorder (MDD). There is a clear need for a thorough discussion of the methodological issues concerning the use of these scales, as suggested also by Demyttenaere and De Fruyt in a recent review [1], because clinical trials researchers in depression continue to struggle with the ability to detect signals of the efficacy of antidepressant agents.

**Keywords** Depression · Rating scales · Questionnaires · Assessment · Mood · Psychiatry

### Gold Standard Rating Scales

- The Hamilton Rating Scale for Depression (HAM-D or HRSD) [2] (clinician-administered)
- The Beck Depression Inventory (BDI) [27] (patient-rated)
- Inventory of Depressive Symptomatology (IDS or QIDS) [35] (patient-rated or clinician administered)

---

C. Cusin (✉)

Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Wang Ambulatory Care Center 812, 15 Parkman Street, Boston, MA 02114, USA  
e-mail: ccusin@partners.org

## ***Hamilton Rating Scale for Depression (HAM-D or HRSD)***

This is one of the earliest scales to be developed for depression, and is a clinician-rated scale aimed at assessing depression severity among patients. The original HAM-D included 21 items, but Hamilton pointed out that the last four items (diurnal variation, depersonalization/derealization, paranoid symptoms, and obsessive-compulsive symptoms) should not be counted toward the total score because these symptoms are either uncommon or do not reflect depression severity [2].

Therefore, the 17-item version of the HAM-D (reproduced in the appendix to this chapter) has become the standard for clinical trials and, over the years, the most widely used scale for controlled clinical trials in depression (we found in a recent Medline search that more than 500 studies have used the HAM-D as primary efficacy measure). Its widespread use, however, has not prevented investigators from recognizing the limitations of this instrument and from trying to improve it. The main limitations of the original 17-item version of the HAM-D were recognized to be (1) the failure to include all symptom domains of major depressive disorder (MDD), in particular, reverse neurovegetative symptoms, (2) the presence of items measuring different constructs (e.g., irritability and anxiety, loss of interest and hopelessness), and (3) the uneven weight attributed to different symptom domains (e.g., insomnia may be rated up to 6 points, while fatigue only up to 2).

### **Application of Scale**

#### Method of Administration

The scale is widely used in clinical trials and in clinical practice, and in general is administered weekly. To improve inter-rater reliability, a structured interview guide for the HAM-D was developed in 1988 by Janet Williams (SIGHD) [3, 4] and her guide soon became the gold standard for training and for clinical studies (see for example: <http://www.ids-qids.org/translations/english/SIGHD-IDSEnglish-USA.pdf>). We recommend using the interview guide to improve inter-rater reliability.

#### Timing of Administration

Considering the busy schedule of both patients and health professionals, the time needed to administer a scale could represent a significant burden. Our research has found that the average duration of the HAM-D interviews was 12 minutes. However, our estimations of the length of those interviews are underestimates and features of depression such as psychomotor retardation may significantly increase their duration. It is noteworthy that in our simulation, using a structured interview did not seem to considerably increase the duration of the administration of the scale.

Because of its widespread use over the course of decades, the HAM-D is the most popular depression severity measure in the history of MDD trials, and is very familiar to most clinical researchers in the area of depression.

### Reliability and Internal Consistency

The HAM-D is a multidimensional scale, and this implies that the score of a specific item cannot be considered a good predictor of the total score [5]. It also means that identical total scores from two different patients may have different clinical meanings (i.e., a very high rating on few items can yield the same score as a moderate rating on many items) [6]. A number of studies have shown the internal consistency of different versions of HAM-D to range widely from 0.48 to 0.92. Higher coefficient alpha values were reached with the use of a structured interview (see [7] for more details). A recent study reported internal consistency coefficients of 0.83 for HAM-D-17 and 0.88 for HAM-D-24 [8]. A complete review of the psychometric properties of the HAM-D has been published recently. In this paper, the authors reviewed 70 studies on psychometric properties of the HAM-D, published since 1979, and showed that the majority of HAM-D items have adequate reliability [9].

### Inter-rater Reliability

Inter-rater reliability has been reported to be very high for HAM-D total scores (0.80–0.98), even if it is poor for some of its items. All items showed adequate reliability when the scale was administered with interview guidelines [10]. A sufficiently high inter-rater reliability ( $>0.60$ ) was reported for most of the HAM-D items and the total score (0.57–0.73) in a study on inter-rater reliability in 21 psychiatric novices who had negligible previous experience with the HAM-D [11]. This score appears to be improved greatly with the use of appropriate training and structured interview [12].

### Test–Retest Reliability

Test–retest reliability for the HAM-D using the Structured Interview Guide has been reported to be as high as 0.81, even among minimally trained raters from multiple disciplines [4, 13, 14].

### Validity

Validity of the HAM-D has been reported to range from 0.65 to 0.90 with global measures of depression severity, and to be highly correlated with clinician-rated measures such as MADRS and IDS-C [7].

### Scoring Key

The total score is obtained by summing the score of each item, 0–4 (symptom is absent, mild, moderate, or severe) or 0–2 (absent, slight or trivial, clearly present). For the 17-item version, scores can range from 0 to 54.

## Cut-Off Scores

It is accepted by most clinicians that scores between 0 and 6 do not indicate the presence of depression, scores between 7 and 17 indicate mild depression, scores between 18 and 24 indicate moderate depression, and scores over 24 indicate severe depression. A total HAM-D score of 7 or less after treatment is for most raters a typical indicator of remission [15]. A decrease of 50% or more from baseline during the course of the treatment is considered indicator of clinical response, or in other words, a clinically significant change.

## ***Beck Depression Inventory (BDI)***

The gold standard of self-rating scales is the Beck Depression Inventory (BDI) [16], which was initially developed to assess the efficacy of psychoanalytically oriented psychotherapy in depressed subjects. The BDI is copyrighted by Harcourt Assessment, Inc., and so is not reproduced in this chapter. Information about purchase of this scale and manual are available from their website at: <http://harcourtassessment.com/haiweb/cultures/en-us/productdetail.htm?pid=015-8018-370>.

This scale was designed to measure the severity of depressive symptoms that the test taker is experiencing “at that moment.” The original BDI included 21 items concerning different symptom domains, with four possible answers describing symptoms of increasing severity associated with a score from 0 to 3. It was later amended to BDI-IA [17], and after the publication of the DSM-IV, to the BDI-second edition (BDI-II) [18]. Four new items (agitation, worthlessness, concentration difficulty, and loss of energy) were added to make the BDI-II more reflective of DSM-IV criteria of MDD, and some BDI-IA items (i.e., weight loss, body image change, work difficulty, and somatic preoccupation) were eliminated because they were considered less indicative of the overall severity of depression. Beck and colleagues also rewrote almost all other BDI-II items for clarity, and the time frame for ratings was extended from 1 to 2 weeks [19, 20].

Self-rating scales, such as the BDI, offer some advantages over clinician-rated scales, as they may take less time, do not require trained personnel, and their administration and scoring process appear more standardized [21]. Self-rating scales also require that individuals are able to read at a minimal reading level, and that they speak the language used in at least one translation of the scale.

## **Reliability and Validity**

### Reliability

#### *Internal Consistency*

Beck and colleagues in 1988 published a meta-analysis of all the psychometric studies on the BDI from 1961 to June 1986 and found a mean coefficient alpha of 0.86 for psychiatric subjects [22]. In 1996, after the publication of the BDI-II, Beck and

coworkers compared the BDI-II and BDI-IA scales in a sample of 140 psychiatric outpatients with various psychiatric disorders and found coefficient alpha for the BDI-II and the BDI-IA of 0.91 and 0.89, respectively [19]. The BDI and the BDI-II were also tested on a larger sample ( $n = 500$ ), where the BDI-II showed improved clinical sensitivity, with reliability ( $\alpha = 0.92$ ) higher than the BDI ( $\alpha = 0.86$ ) (Psychological Corporation Website, 2003).

### *Test–Retest Reliability*

With self-administered measures, assessing test–retest reliability may be complicated by the fact that the correlation coefficient may increase spuriously because of practice or because of memory effects. However, in a Spanish study, test–retest reliability for the BDI was between 0.65 and 0.72 [23].

### *Validity*

The convergent validity with the BDI has been reported to be extremely variable, ranging between 0.27 and 0.89 [24]. Beck and colleagues showed that in psychiatric patients, the mean correlations of the BDI were 0.72 with clinical ratings and 0.73 with the HAM-D [22] and 0.57–0.83 with the Zung SDS [25].

## ***Inventory of Depressive Symptomatology***

In the 1980s, John Rush and colleagues [35] developed and published the clinician-rated Inventory of Depressive Symptomatology (IDS) (reproduced in the appendix to this chapter), which was intended to “remedy the deficits of the HAM-D and the MADRS” by including all the symptom domains of the DSM-based MDD, as well as both melancholic and atypical (e.g., reversed neurovegetative) features, by scaling each item to allow for the measurement of milder forms of MDD, providing clearer items definition (for example, irritability and anxiety were rated separately) and equivalent weight for each symptom domain. The original IDS had 28 items [35], while an additional two items (leaden paralysis; interpersonal rejection sensitivity) were added later to better capture atypical MDD features [36]. Subsequently, Rush and colleagues selected 16 items from the IDS-30, assessing the DSM-IV diagnostic criteria for MDD, and assembled them in the short version of the IDS, namely the Quick Inventory of Depressive Symptomatology (QIDS) [8]. Dr. Rush and Colleagues created a self-rated version of the 28-item IDS-C in the 1980s, called the IDS-SR-28 [35, 37], then added the two items of atypical MDD features to obtain the 30-item version [36], and shortened it to the 16 items of the DSM-IV diagnostic criteria for the QIDS-SR [8] (reproduced in the appendix to this chapter).

### **Scoring Key**

For all the versions, add the scores of the items to obtain the total score, except for items 11–12 (increased or decreased appetite) and 13–14 (increased or decreased

weight) for which the highest of the two has to be included. A description of cut-offs for moderate and severe depression for the different versions is available at the website <http://www.ids-qids.org/index2.html#table2>.

## **Reliability**

### **Internal Consistency**

Internal consistency of the IDS is high. In a study published in 1999 on 68 patients assessed at admission, after 5, 10, and 28 days of antidepressant treatment, the Cronbach's alpha coefficients reported were 0.75 for the IDS-C and 0.79 for the IDS-SR [38]. Alpha values were reported to vary from 0.67 to 0.82 for subjects with current depression in a very large sample [36]. In another study on 544 outpatients with MDD and 402 outpatients with bipolar disorder, the Cronbach's alpha ranged from 0.81 to 0.94 for all four scales (QIDS-C16, QIDS-SR16, IDS-C30, and IDS-SR30) [39]. Cronbach's alpha ranged from 0.81 to 0.90 for the QIDS-C and was reported to be 0.86 for the QIDS-SR ([http://www.ids-qids.org/IDS\\_Website\\_Document.pdf](http://www.ids-qids.org/IDS_Website_Document.pdf)).

### **Inter-rater Reliability**

Inter-rater reliability for the IDS-C was reported to be very high (0.96). ([http://www.ids-qids.org/IDS\\_Website\\_Document.pdf](http://www.ids-qids.org/IDS_Website_Document.pdf)).

## **Validity**

IDS-SR correlation with the HAM-D-24 and BDI have been investigated in a sample of 289 patients with mixed diagnoses and reported to be respectively 0.67 and 0.78, while the IDS-C was highly correlated with the HAM-D ( $r = 0.92$ ) and less with the BDI ( $r = 0.61$ ) in a sample of 82 outpatients [35]. In another very large sample ( $n = 596$ ) of patients treated for chronic non-psychotic MDD, the QIDS-SR total scores were highly correlated with IDS-SR-30 (0.96) and with the HAM-D-24 (0.86) total scores [8]. The QIDS-C and QIDS-SR scores have been reported to be correlated (0.72 or more) with those of the HAM-D-17 ([http://www.ids-qids.org/IDS\\_Website\\_Document.pdf](http://www.ids-qids.org/IDS_Website_Document.pdf)) and HAM-D-24 [40].

## **Other Scales Available for Rating Depression**

### ***Montgomery–Asberg Depression Rating Scale***

The clinician-rated Montgomery and Asberg Depression Rating Scale (MADRS) [reproduced in the Appendix to this chapter] was developed in the late 1970s [26] and this 10-item scale was designed to be sensitive to the effects of antidepressant medications, primarily tricyclic antidepressants (TCAs) [26]. Because this scale was

never updated or modified, it does not target reverse neurovegetative symptoms. It is commonly used in clinical studies and in clinical practice, administered weekly. Structured interview guides for the MADRS have been developed by a number of investigators [13, 27–29].

## **Reliability**

### Internal Consistency

The MADRS appears to be a unidimensional scale, more focused toward psychological, as opposed to somatic aspects of depression [30]. The internal consistency of the MADRS is considered very high, given the high correlation between all items ( $r = 0.95$ ) [31]. In a recent psychometric re-analysis of primary efficacy measures derived from a trial on citalopram efficacy in maintenance therapy of elderly depressed patients, the internal consistency of the MADRS, was found to be superior to that of the HAM-D-17 [6].

### Inter-rater Reliability

One of the original goals of the MADRS was to obtain an instrument that could be used by both psychiatrists and professionals without a specific or with minimal psychiatric training. From the original report of the MADRS, the inter-rater reliability ranged from 0.89 to 0.97 [26]. However, in a German study, significant differences resulted when the same patient was rated by various groups of caregivers (psychiatrists, psychologists, students, and psychiatric nurses) [32].

### Validity

Correlation of MADRS has been shown to be generally high or very high with the HAM-D (between 0.80 and 0.90) [7, 33], RDC (0.70) [34], and with IDS-C (0.81) [34].

### Cut-Off Scores

A score greater than 30 or 35 on the MADRS indicates severe depression, while a score of 10 or below indicates remission.

## ***Zung Self-Report Depression Scale***

The Zung Self-Report Depression Scale (Zung SDS) [41] (reproduced in the appendix to this chapter) was published a few years later than the BDI. It is a 20-item self-report index that covers, in varying degree, a broader spectrum of symptoms than the BDI, including psychological, affective, cognitive, behavioral, and somatic aspects of depression.

## Scoring Key

Respondents are instructed to rate each item on a scale ranging from 0 to 4 in terms of “how frequently” they have experienced each symptom, instead of “how severe.” The time frame was originally “at the present,” but in subsequent version the time frame was extended to one week, therefore recommending weekly administration. A total score is derived by summing the individual item scores (1–4), and ranges from 20 to 80. The items are scored as follows: 1 = a little of the time, through 4 = most of the time, *except* for items 2, 5, 6, 11, 12, 14, 16, 17, 18, and 20 which are scored inversely (4 = a little of the time).

## Cut-Off Scores

Most people with depression score between 50 and 69, while a score of 70 and above indicates severe depression. No revision of the scale was made after the original publication and is nowadays less used in clinical practice.

## Validity

The correlation between Zung SDS and HAM-D was reported to range between 0.68 and 0.76, being lower with HAM-D at baseline [21]. The best results were observed at mild or moderate severity levels, while the greatest disagreement between Zung and HAM-D was observed for patients with non-endogenous symptom patterns [42].

## Other Issues in Assessing Depression

### *Ability of Depression Rating Scales to Detect Clinical Changes with Treatment*

The ability of psychometric instruments to detect changes related to treatment is a concept that has been extensively discussed by Robert Kellner [43]. In his review of the literature, he indicated the importance for a measure of capturing changes over time, particularly in those symptoms characterizing MDD [43]. As Kellner stated, a scale may be valid but have low sensitivity to detect change in the state of the patient. For example, a scale may contain items relatively insensitive to change and therefore may be highly stable and underestimate the effects of a treatment. The BDI measures attitudes and cognitions which are fairly stable over time among depressed patients, and therefore may underestimate the degree of improvement during acute pharmacological treatments. In addition, a scale might have items accurately measuring mild depression, but may be less sensitive to moderate or severe depression, leading to a poor sensitivity to detect improvements in patients with more severe depression at baseline. The scales actually used in clinical trials typically are considered to have a relatively good sensitivity to change, with the exception of the Zung scale, which is considered more sensitive to differences across subgroups of patients, than to change over time [44].

### ***Minimizing Biases in the Assessment of Depression Symptom Domains***

A possible bias in measurement of depressive symptoms may be related to the variable emphasis on somatic versus psychological symptoms. For example, since 3 of the 17 items of the HAM-D concern sleep disturbances (insomnia) and contribute up to 11.5% of the total score, it has been hypothesized that the HAM-D may favor sedating antidepressant drugs (i.e., some TCAs or trazodone), which may improve sleep, regardless of “true” antidepressant effects. Similarly, drugs associated with side effects such as sleep disturbances, gastrointestinal (GI) symptoms, agitation, and nervousness, such as the SSRIs and the SNRIs, could be associated with an artificially elevated HAM-D score at endpoint, thereby underestimating improvement.

When considering somatic symptoms, the convention is often that such symptoms should be rated at face value, without trying to distinguish side effects from symptoms. This approach may affect all measures of depression severity, as sleep and appetite disturbances may be side effects and/or symptoms of MDD. However, in the case of the HAM-D, psychological, and somatic symptoms/side effects such as anxiety/agitation, sexual dysfunction, dry mouth, and diarrhea may be affecting the score to a greater degree than other scales [45]. The BDI, MADRS, HAM-D-6, IDS, and QIDS are considered to be relatively insensitive to this well-known bias of the HAM-D [46].

### ***Ability of Depression Rating Scales to Measure Symptoms Across Depressive Subtypes***

Since major depressive disorder is not a homogeneous clinical entity, a valid scale must measure symptoms across all subtypes, allowing clinicians to compare treatment efficacy in various depressive populations. In fact, inaccurate assessments across subtypes have been hypothesized to be one of the culprits for the high failure rate of many MDD clinical trials [46, 47]. Due to the differences in historical background and rationale behind each rating scale, the HAM-D, the MADRS, and the IDS/QIDS have different levels of ability to reflect the heterogeneity of MDD and to capture symptoms characteristic of depressive subtypes. The HAM-D-28, the IDS, and the BDI-II cover symptoms of both atypical and melancholic depression, while atypical symptoms are far less relevant in the BDI and the Zung scale, where they represent only 5% of the total score, and in the MADRS where these symptoms are not included at all.

### ***Self- Versus Clinician-Administered Depression Rating Scales***

The dilemma between self-administered and clinician-rated scales has led to a number of studies investigating differences and similarities between those two ways of assessing depressive symptoms. Although concordance rates between self ratings and observer ratings are generally acceptable, significantly discordant ratings

have been obtained in many studies showing that clinicians and patients rate the depressive symptoms differently [48–50]. Clinicians are thought to measure depressive severity more accurately [37, 51]. In fact, in a study of the two versions of IDS (IDS-C and IDS-SR), where these two scales were administered to 64 inpatients with MDD on day 1, 10, and day 28 after antidepressant treatment, the self-rated version of IDS showed a lesser sensitivity to change over time compared to the clinician-rated version [38]. On the other hand, self-rated scales may be more sensitive to detect changes than clinician-rated scales in milder forms of depression. In fact, a study compared the scores from three different scales, HAM-D, IDS-C, and IDS-SR, across severity subgroups in patients with dysthymic depression, non-endogenous MDD, and endogenous MDD. More symptoms were self-reported by the dysthymic patients and the non-endogenous patients than recorded by the clinician, but for the endogenously depressed patients self-reported and clinician-rated symptoms were comparable [37]. Similarly, a study published in 2000 showed that the discrepancies between BDI and HAM-D-21 scores were increased in patients with younger age, higher educational level, atypical depressive subtype, and neurotic personality features, all those factors being associated with higher BDI scores [52]. Sayer et al. [53] investigated the correlation between the HAM-D-24 and the BDI in 114 severely depressed inpatients, treated with electroconvulsive therapy. Their study showed a relatively poor correlation between the instruments at baseline, due to a specific subgroup of depressive patients who were evaluated by the observer as severely depressed, but rated themselves as less symptomatic. Some clinical features of the subgroup were advanced age, less education, presence of psychosis, lack of insight, and severe hypochondriasis. This same subgroup showed the greatest improvement in HAM-D score and contributed largely to the discrepancy in effect size between HDRS and BDI ratings.

When the effect sizes (calculated as the difference between the proportions of responders taking drug and those taking placebo) derived from patient self-ratings and from clinician ratings were compared by Petkova and colleagues, the result was that the self-rating scales were associated with smaller effect sizes, therefore supporting the hypothesis that they are less likely to differentiate active drug from placebo [54]. However, the self-rating scales in the Petkova study did not include scales, such as the IDS-SR or QIDS-SR, which are reported to show more robust performance in clinical trials compared to the older self-rating scales.

In clinical practice, different clinicians choose what scale to administer according to their level of comfort with a scale and to the time available. Some choose to present self-rating scales (most often used are the BDI, IDS-SR or QIDS-SR) to patients in the waiting room and have them fill out the questionnaires. Other clinicians prefer asking patients directly about symptoms and administer the scale themselves during the visit (HAM-D, MADRS or IDS-C), in particular with complicated patients or patients with comorbidities for which answers about physical symptoms may need clarification. The clinician should be aware of strengths and limitations of at least few of the most commonly used scales, and should be able to choose the most appropriate instrument for the patient.

## *Assessing Depression Across Age Groups*

Depression is very common among elderly patients, whose depressive psychopathology has been shown to be different in some aspects from younger individuals, i.e., increased prevalence of sleep disturbances and hypochondriasis [55]. Elderly depressed are more likely to be affected by medical conditions that complicate their evaluation and their treatment. For example, the presence of somatic symptoms due to concomitant medical illnesses may be misattributed to the depression or vice versa [56]. Linden et al. [57] reported that in depressed patients who were 70 years or older and also suffered from a medical illness, eight items of the HAM-D may be elevated by the concurrent somatic disorder (somatic anxiety, GI symptoms, general somatic symptoms, hypochondriasis, weight loss, middle insomnia, and work). In other cases, older patients with clinically significant depression may underreport their symptoms [58]. In addition, the presence of cognitive symptoms may impair the evaluation of depression, as they might be related to natural cognitive functioning decline, to the onset of dementing disorders, or to depression itself. Nebes et al. [59] measured the working memory, information-processing speed, episodic memory, and attention over a 12-week randomized, double-blind trial with nortriptyline and paroxetine. Compared to the elderly controls, cognitive dysfunction persisted in older depressed patients, even after their depression had responded to antidepressant medications. Cognitive symptoms may affect patients' ability to understand and/or respond appropriately to questions about their depressive symptoms. Finally, items assessing thoughts of death, pessimism, and reduced interest or activity may have a different meaning in a geriatric population compared to younger adults. Scales have been developed with the specific purpose of screening for MDD in the geriatric population, of which the gold standard is the Geriatric Depression Scale (GDS), a self-report scale with different versions containing 30, 15, and 4 items [60, 61]. Other scales are the Brief Assessment Schedule Depression Cards (BASDEC), the Cornell Scale for Depression in Dementia and the Geriatric Mental State Schedule (GMSS) (for a review see [62]). Despite the differences in symptoms between geriatric and adult patients with MDD, the primary outcome measures used for the antidepressant trials in the elderly (age  $\geq 65$  years) are still the scales developed in the adult population such as the HAM-D, the BDI and the MADRS [63–66]. However, further studies are necessary to compare the performance of different scales in this specific population.

Similarly, depressive symptoms may be different in children and adolescents from those of adults, challenging the use in children of scales aimed to assess depression among adults. In addition, scales used for adults often use anchor points that are best suited to capture symptoms in adult populations, and may be less useful for children and adolescents. Furthermore, as Poznanski pointed out, the measure of the non-verbal behavior for children and adolescents was most strongly associated with the diagnosis of depression and was also the best predictor of the severity of depression [67]. Many authors have tried to develop instruments to measure depression in children and adolescents. The Children's Depression Rating Scale (CDRS) and its revised version (CDRS-R) are clinician-rated instruments to measure severity

of depression in children [67, 68]. The CDRS has been validated for use in children and adolescents [68] and has been used as a primary outcome measure in clinical trials [70, 71].

Self-rated scales are also commonly used in children and adolescents, such as the Kutcher Adolescent Depression Scale (KADS), the Children's Depression Inventory (CDI) [72], the Child Depression Scale [73–75], and the Beck Youth Inventories of Emotional and Social Impairment [76]. Brooks et al. suggested that the 11-item KADS is a sensitive measure of treatment outcome in adolescents diagnosed with MDD [77].

### *Assessing Depression Across Different Cultures*

Cross-cultural variations in presenting symptoms of depression have been reported [78]. For example, certain symptoms, such as self-blame and guilt, are not common to all cultures [79, 80]. In addition, differences have been observed in the severity of decision-making impairments in depression across cultures [81]. Researchers from our group have also observed higher rates of suicidal ideation among Asian-Americans (24%), participants who report ethnic heritage as “Other” (19.5%) Caucasians (16.9%), and Asian-Indians (14%), compared to Hispanics (7.3%) and African-Americans (6%) in a sample of 707 college students [82]. Psychotic symptoms have also been found to be more prevalent in Hispanic patients with MDD seeking treatment, compared to Caucasians and Portuguese patients, but not when compared to African-American [83].

The most striking and consistent finding of cross-cultural studies on depression is the variation in the somatization domain. After screening approximately 26,000 patients for MDD at 15 primary care centers in 14 countries and 5 continents, Simon and colleagues found that the prevalence of somatic symptoms varied across centers from 45% to 95% [84]. Moreover, not only the frequency, but also the type of somatic complaints may be subject to cultural influences, as shown in a study on inpatients admitted for MDD in Greece ( $N = 60$ ) and in Australia ( $N = 56$ ) [85]. Higher rates of somatization have been also reported in depressed Japanese, Chinese, and Turkish patients compared to their western counterparts diagnosed with MDD [86–88]. Relevant differences have also been observed in self-reported scales. Fugita and colleagues analyzed the Zung SDS scores in students from four different countries. Korean and Philippine students had the highest scores, Caucasian Americans the lowest [89]. The relatively greater depression severity in Asian-American populations was confirmed in a recent study comparing BDI results between a sample of Asian-American ( $n = 238$ ) and Caucasian-American students ( $n = 556$ ) [90]. Cross-cultural comparison studies have typically not used outcome measures such as the HAM-D, the MADRS, the IDS and the QIDS, even though all have translated versions available in more than 20 languages. Because of cross-cultural and cross-ethnic differences in patients with MDD, one may argue that scales that were developed for the assessment of depression among Western European and North American Caucasians may not be culturally sensitive in measuring

symptoms across other ethnic and cultural groups. However, there is no good evidence that these scales fail to perform well in clinical trials conducted in different countries.

### ***Assessing Depression Across Different Educational and Comprehension Levels***

To effectively assess severity of depressive symptoms through a clinician-administered questionnaire, it is necessary that patients understand the meaning of the questions asked. Although readability is widely used as a proxy for comprehension, it might give a false sense of confidence about comprehensibility. In fact, when respondents lacked not only the cognitive capacity to fully understand a standardized question, but also the motivation to answer it thoughtfully, patients often produce a superficially adequate answer (i.e., choosing the first or last response, choosing a neutral response, choosing a socially desirable response or repeating the previous response) [91]. Finally in situations in which respondents' motivation and/or time are limited, even individuals who could understand a complex instrument may not make the effort to answer questions thoughtfully [92].

### ***Assessing Depression with Psychiatric Comorbidities***

Little is known about the ability of scales to measure changes in depressive symptomatology across populations with varying degrees of psychiatric comorbidity. For example, it is well known that comorbid anxiety disorders are very common in MDD and the presence of a comorbid anxiety disorder can influence the anxiety and somatic items and therefore inflate the total score of a multidimensional scale such as the HAM-D. Furthermore, core obsessive-compulsive disorder (OCD) symptoms may heavily affect ratings on items covering guilt feelings (because of aggressive/sexual obsessions), work and activities (reduced if the patients are immersed in their compulsions), and anxiety [92]. When a comorbid eating disorder is not an exclusion criterion, the relative influence of items related to weight change, irregular eating habits, guilt, and GI and somatic symptoms has to be carefully considered. For example in the HAM-D 17, the sum of items covering feeling guilty, weight change, somatic anxiety, and gastrointestinal symptoms, may represent 33.6% of the total score, but only 22.2% and 20% of the QIDS and MADRS scores, respectively.

### ***Assessing Depression with Medical Comorbidities***

Assessment of depression in medically ill populations is complicated by the fact that emotional, behavioral, or cognitive symptoms may be caused by the concomitant

medical illness and/or by the medications used to treat the illness. Ideally, depression assessments should be restricted to variables and items that avoid confounding by medical illness. Two measures have been designed for assessing depression in the medical patients by excluding somatic items: the Hospital Anxiety Depression Scale (HADS) [93] and the Beck Depression Inventory for Primary Care (BDI-PC); however, most of the depression measures developed for medically ill populations have not been adequately tested as outcome measure in depression trials.

**Acknowledgments** This work has been funded in whole or in part with Federal funds from the National Institute of Mental Health, National Institutes of Health, under Contract N01MH90003 (STAR\*D) and GMO N01MH90003 (NIMH Depression Trials Network). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

The authors thank Dr Janet B.W. Williams, Dr Kenneth Kobak, and Dr John Rush for their contributions to this chapter.

## References

1. Demyttenaere K, De Fruyt J: Getting what you ask for: on the selectivity of depression rating scales. *Psychother Psychosom* 2003; 72:61–70
2. Hamilton M: A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; 23:56–62
3. Williams JBW: Research assessments: to standardize or not to standardize?, in *Research Designs and Methods in Psychiatry*. Edited by Fava M, Rosenbaum JF, Amsterdam, Elsevier, 1992, pp 31–36
4. Williams JB: A structured interview guide for the Hamilton depression rating scale. *Arch Gen Psychiatry* 1988; 45:742–7
5. Bech P: Measurement issues., in *Textbook of biological psychiatry. Part 1: Basic principles*. Edited by D'Haenen H. New York, Wiley, 2002
6. Bech P, Tanghøj P, Andersen HF, Overo K: Citalopram dose-response revisited using an alternative psychometric approach to evaluate clinical effects of four fixed citalopram doses compared to placebo in patients with major depression. *Psychopharmacology (Berl)* 2002; 163:20–5
7. Hamilton M: Hamilton rating scale for Depression (Ham-D), in *Handbook of psychiatric measures*. Washington DC, APA, 2000, pp 526–8
8. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, Markowitz JC, Ninan PT, Kornstein S, Manber R, Thase ME, Kocsis JH, Keller MB: The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 2003; 54:573–83
9. Bagby RM, Ryder AG, Schuller DR, Marshall MB: The Hamilton depression rating scale: has the gold standard become a lead weight? *Am J Psychiatry* 2004; 161:2163–77
10. Moberg PJ, Lazarus LW, Mesholam RI, Bilker W, Chuy IL, Neyman I, Markvart V: Comparison of the standard and structured interview guide for the Hamilton Depression Rating Scale in depressed geriatric inpatients. *Am J Geriatr Psychiatry* 2001; 9:35–40
11. Muller MJ, Dragicevic A: Standardized rater training for the Hamilton Depression Rating Scale (HAM-D-17) in psychiatric novices. *J Affective Dis* 2003; 77:65–9
12. Kobak KA, Lipsitz JD, Feiger A: Development of a standardized training program for the Hamilton Depression Scale using internet-based technologies: results from a pilot study. *J Psychiatric Res* 2003; 37:509–15

13. Takahashi N, Tomita K, Higuchi T, Inada T: The inter-rater reliability of the Japanese version of the Montgomery-Asberg depression rating scale (MADRS) using a structured interview guide for MADRS (SIGMA). *Hum Psychopharmacol* 2004; 19:187–92
14. Davidson J, Turnbull CD, Strickland R, Miller R, Graves K: The Montgomery-Asberg Depression Scale: reliability and validity. *Acta Psychiatrica Scandinavica* 1986; 73:544–8
15. Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, Rush AJ, Weissman MM: Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry* 1991; 48:851–5
16. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J: An inventory for measuring depression. *Arch Gen Psychiatry* 1961; 4:561–71
17. Beck AT, Steer RA: Manual for the Beck Depression Inventory. San Antonio, TX, 1993
18. Beck AT, Steer RA, Brown GK: Manual for the Beck Depression Inventory, 2nd ed. San Antonio, TX, 1996
19. Beck AT, Steer RA, Ball R, Ranieri W: Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *J Pers Assess* 1996; 67:588–97
20. Steer RA, Clark DA, Beck AT, Ranieri WF: Common and specific dimensions of self-reported anxiety and depression: the BDI-II versus the BDI-IA. *Behav Res Ther* 1999; 37:183–90
21. Biggs JT, Wylie LT, Ziegler VE: Validity of the Zung Self-rating Depression Scale. *Br J Psychiatry* 1978; 132:381–5
22. Beck AT, Steer RA, Carbin M: Psychometric properties of the Beck depression inventory: Twenty-five years of evaluation. *Clin Psychol Rev* 1988; 8:77–100
23. Vazquez C, Sanz J: Fiabilidad y validez factorial de la versión española del inventario de depresión de Beck., in III Congreso de Evaluación Psicológica. Barcelona, 1991
24. Richter P, Werner J, Heerlein A, Kraus A, Sauer H: On the validity of the Beck depression inventory. A review. *Psychopathology* 1998; 31:160–8
25. Zung WKK: Zung Self-rating Depression Scale (Zung SDS), in *Handbook of psychiatric measures*. Washington, DC, APA, 2000, pp 534–537
26. Montgomery SA, Asberg M: A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; 134:382–9
27. Lobo A, Chamorro L, Luque A, Dal-Re R, Badia X, Baro E: Validation of the Spanish versions of the Montgomery-Asberg depression and Hamilton anxiety rating scales. *Med Clin (Barc)* 2002; 118:493–9
28. Fleck MP, Guelfi JD, Poirier-Littre MF, Loo H: Application of a structured interview guide adapted to 4 depression scales. *Encephale* 1994; 20:479–86
29. Williams JB: Structured Interview Guide for the Montgomery Asberg Depression Rating Scale (SIG-MA). Biometrics Research Department, 1996
30. Fava M: Somatic symptoms, depression, and antidepressant treatment. *J Clin Psychiatry* 2002; 63:305–7
31. Galinowski A, Lehert P: Structural validity of MADRS during antidepressant treatment. *Int Clin Psychopharmacol* 1995; 10:157–61
32. Schmidtke A, Fleckenstein P, Moises W, Beckmann H: Untersuchungen zur Reliabilität und Validität einer deutschen Version der Montgomery-Asberg Depression-Rating Scale (MADRS). *Schweizer Archiv für Neurologie und Psychiatrie* 1988; 139:51–65
33. Muller MJ, Himmerich H, Kienzle B, Szegedi A: Differentiating moderate and severe depression using the Montgomery-Asberg depression rating scale (MADRS). *J Affective Dis* 2003; 77:255–60
34. Montgomery SA, Asberg M: Montgomery-Asberg depression rating scale, in *Handbook of psychiatric measures*. Washington, DC, APA, 2000, pp 531–33
35. Rush AJ, Giles DE, Schlessler MA, Fulton CL, Weissenburger J, Burns C: The Inventory for Depressive Symptomatology (IDS): preliminary findings. *Psychiatry Res* 1986; 18:65–87
36. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH: The inventory of depressive symptomatology (IDS): psychometric properties. *Psychol Med* 1996; 26:477–86

37. Rush AJ, Hiser W, Giles DE: A comparison of self-reported versus clinician-related symptoms in depression. *J Clin Psychiatry* 1987; 48:246–8
38. Corruble E, Legrand JM, Duret C, Charles G, Guelfi JD: IDS-C and IDS-sr: psychometric properties in depressed in-patients. *J Affect Dis* 1999; 56:95–101
39. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, Crismon ML, Shores-Wilson K, Toprac MG, Dennehy EB, Witte B, Kashner TM: The inventory of depressive symptomatology, clinician rating (IDS-C) and self-report (IDS-SR), and the quick inventory of depressive symptomatology, clinician rating (QIDS-C) and self-report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med* 2004; 34:73–82
40. Rush AJ, Trivedi MH, Carmody TJ, Ibrahim HM, Markowitz JC, Keitner GI, Kornstein SG, Arnow B, Klein DN, Manber R, Dunner DL, Gelenberg AJ, Kocsis JH, Nemeroff CB, Fawcett J, Thase ME, Russell JM, Jody DN, Borian FE, Keller MB: Self-reported depressive symptom measures: sensitivity to detecting change in a randomized, controlled trial of chronically depressed, nonpsychotic outpatients. *Neuropsychopharmacology* 2005; 30:405–16
41. Zung WW: A self-rating depression scale. *Arch Gen Psychiatry* 1965; 12:63–70
42. White J, White K, Razani J: Effects of endogeneity and severity on consistency of standard depression rating scales. *J Clin Psychiatry* 1984; 45:260–1
43. Kellner R: The development of sensitive scales for research in therapeutics, in *Research designs and methods in Psychiatry.*, vol 9. Edited by Fava M, Rosenbaum JF. Amsterdam, The Netherlands, Elsevier, 1992, pp 213–222
44. Moran P, Lambert MJ: Measurement methods in affective disorders: consistency with DSM-III diagnosis., in *The Assessment of Psychotherapy Outcome*. Edited by Lambert MJ, DeJulio SS, Christensen ER. New York, Wiley- Interscience, 1983, pp 263–303
45. Maier W, Philipp M: Improving the assessment of severity of depressive states: a reduction of the Hamilton Depression Rating Scale. *Pharmacopsychiatry* 1985; 18
46. Moller HJ: Methodological aspects in the assessment of severity of depression by the Hamilton Depression Scale. *Eur Arch Psychiatry Clin Neurosci* 2001; 251 Suppl 2:II13–20
47. Williams JB: Standardizing the Hamilton depression rating scale: past, present, and future. *Eur Arch Psychiatry Clin Neurosci* 2001; 251 Suppl 2:II6–12
48. Bailey J, Coppen A: A comparison between the Hamilton Rating Scale and the Beck Inventory in the measurement of depression. *Br J Psychiatry* 1976; 128:486–9
49. Domken M, Scott J, Kelly P: What factors predict discrepancies between self and observer ratings of depression? *J Affective Dis* 1994; 31:253–9
50. Prusoff BA, Klerman GL, Paykel ES: Concordance between clinical assessments and patients' self-report in depression. *Archives of General Psychiatry* 1972; 26:546–52
51. Berrios GE, Chen EY: Recognising psychiatric symptoms. Relevance to the diagnostic process. *Br J Psychiatry* 1993; 163:308–14
52. Enns MW, Larsen DK, Cox BJ: Discrepancies between self and observer ratings of depression. The relationship to demographic, clinical and personality variables. *J Affective Dis* 2000; 60:33–41
53. Sayer NA, Sackeim HA, Moeller JR, Prudic J, Devanand DP, Coleman EA, Kierksky JE: The relations between observer-rating and self-report of depressive symptomatology. *Psychol Assess* 1993; 5:350–360
54. Petkova E, Quitkin FM, McGrath PJ, Stewart JW, Klein DF: A method to quantify rater bias in antidepressant trials. *Neuropsychopharmacology* 2000; 22:559–65
55. Koenig HG, Cohen HJ, Blazer DG, Krishnan KR, Sibert TE: Profile of depressive symptoms in younger and older medical inpatients with major depression. *J Am Geriatr Soc* 1993; 41:1169–76
56. Kurlowicz LH, Streim JE: Measuring depression in hospitalized, medically ill, older adults. *Arch Psychiatr Nurs* 1998; 12:209–18
57. Linden M, Borchelt M, Barnow S, Geiselman B: The impact of somatic morbidity on the Hamilton depression rating scale in the very old. *Acta Psychiatrica Scandinavica* 1995; 92:150–4

58. Lyness JM, Cox C, Curry J, Conwell Y, King DA, Caine ED: Older age and the underreporting of depressive symptoms. *J Am Geriatr Soc* 1995; 43:216–21
59. Nebes RD, Pollock BG, Houck PR, Butters MA, Mulsant BH, Zmuda MD, Reynolds CF, 3rd: Persistence of cognitive impairment in geriatric patients following antidepressant treatment: a randomized, double-blind clinical trial with nortriptyline and paroxetine. *J Psychiatr Res* 2003; 37:99–108
60. Yesavage JA: Geriatric Depression Scale. *Psychopharmacol Bull* 1988; 24:709–11
61. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO: Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatric Res* 1982; 17:37–49
62. Burns A, Lawlor B, Craig S: Rating scales in old age psychiatry. *Br J Psychiatry* 2002; 180:161–7
63. Bent-Hansen J, Lunde M, Klysner R, Andersen M, Tanghoj P, Solstad K, Bech P: The validity of the depression rating scales in discriminating between citalopram and placebo in depression recurrence in the maintenance therapy of elderly unipolar patients with major depression. *Pharmacopsychiatry* 2003; 36:313–6
64. Wohlreich MM, Mallinckrodt CH, Watkin JG, Hay DP: Duloxetine for the long-term treatment of major depressive disorder in patients aged 65 and older: an open-label study. *BMC Geriatr* 2004; 4:11
65. Roose SP, Sackeim HA, Krishnan KR, Pollock BG, Alexopoulos G, Lavretsky H, Katz IR, Hakkarainen H: Antidepressant pharmacotherapy in the treatment of depression in the very old: a randomized, placebo-controlled trial. *Am J Psychiatry* 2004; 161:2050–9
66. Sheikh JI, Cassidy EL, Doraiswamy PM, Salomon RM, Hornig M, Holland PJ, Mandel FS, Clary CM, Burt T: Efficacy, safety, and tolerability of sertraline in patients with late-life depression and comorbid medical illness. *J Am Geriatr Soc* 2004; 52:86–92
67. Poznanski E, Mokros HB, Grossman J, Freeman LN: Diagnostic criteria in childhood depression. *Am J Psychiatry* 1985; 142:1168–73
68. Poznanski EO, Cook SC, Carroll BJ: A depression rating scale for children. *Pediatrics* 1979; 64:442–50
69. Poznanski EO, Grossman JA, Buchsbaum Y, Banegas M, Freeman L, Gibbons R: Preliminary studies of the reliability and validity of the children's depression rating scale. *J Am Acad Child Psychiatry* 1984; 23:191–7
70. Emslie GJ, Rush AJ, Weinberg WA, Kowatch RA, Hughes CW, Carmody T, Rintelmann J: A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Arch Gen Psychiatry* 1997; 54:1031–7
71. March J, Silva S, Petrycki S, Curry J, Wells K, Fairbank J, Burns B, Domino M, McNulty S, Vitiello B, Severe J: Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: treatment for adolescents with depression study (TADS) randomized controlled trial. *JAMA* 2004; 292:807–20
72. Kovacs M: The children's depression, inventory (CDI). *Psychopharmacol Bull* 1985; 21:995–8
73. Reynolds WM, Graves A: Reliability of children's reports of depressive symptomatology. *J Abnorm Child Psychol* 1989; 17:647–55
74. Smucker MR, Craighead WE, Craighead LW, Green BJ: Normative and reliability data for the children's depression inventory. *J Abnorm Child Psychol* 1986; 14:25–39
75. Reynolds WM, Anderson G, Bartell N: Measuring depression in children: a multimethod assessment investigation. *J Abnorm Child Psychol* 1985; 13:513–26
76. Steer RA, Kumar G, Beck JS, Beck AT: Evidence for the construct validities of the Beck Youth Inventories with child psychiatric outpatients. *Psychol Rep* 2001; 89:559–65
77. Brooks SJ, Krulewicz SP, Kutcher S: The Kutcher adolescent depression scale: assessment of its evaluative properties over the course of an 8-week pediatric pharmacotherapy trial. *J Child Adolesc Psychopharmacol* 2003; 13:337–49

78. Cheng AT: Case definition and culture: are people all the same? *Br J Psychiatry* 2001; 179:1–3
79. Marsella AJ: Thoughts on cross-cultural studies on the epidemiology of depression. *Cult Med Psychiatry* 1978; 2:343–57
80. Stompe T, Ortwein-Swoboda G, Chaudhry HR, Friedmann A, Wenzel T, Schanda H: Guilt and depression: a cross-cultural comparative study. *Psychopathology* 2001; 34:289–98
81. Radford MH, Nakane Y, Ohta Y, Mann L, Kalucy RS: Decision making in clinically depressed patients. A transcultural social psychological study. *J Nerv Ment Dis* 1991; 179:711–9
82. Sonawalla SB, Kelly KE, Neault NB, Mischoulon D, Farabaugh AH, Pava JA, Yeung A, Fava M: Predictors of suicidal ideation in a college population., in 154th Annual Meeting of the American Psychiatric Association. New Orleans, LA, 2001
83. Posternak MA, Zimmerman M: Elevated rates of psychosis among treatment-seeking Hispanic patients with major depression. *J Nerv Ment Dis* 2005; 193:66–9
84. Simon GE, VonKorff M, Piccinelli M, Fullerton C, Ormel J: An international study of the relation between somatic symptoms and depression. *N Engl J Med* 1999; 341:1329–35
85. Marmanidis H, Holme G, Hafner RJ: Depression and somatic symptoms: a cross-cultural study. *Aust N Z J Psychiatry* 1994; 28:274–8
86. Waza K, Graham AV, Zyzanski SJ, Inoue K: Comparison of symptoms in Japanese and American depressed primary care patients. *Fam Pract* 1999; 16:528–33
87. Parker G, Cheah YC, Roy K: Do the Chinese somatize depression? A cross-cultural study. *Soc Psychiatry Psychiatr Epidemiol* 2001; 36:287–93
88. Ulusahin A, Basoglu M, Paykel ES: A cross-cultural comparative study of depressive symptoms in British and Turkish clinical samples. *Soc Psychiatry Psychiatric Epidemiol* 1994; 29:31–9
89. Fugita SS, Crittenden KS: Towards culture- and population-specific norms for self-reported depressive symptomatology. *Int J Soc Psychiatry* 1990; 36:83–92
90. Lam CY, Pepper CM, Ryabchenko KA: Case identification of mood disorders in Asian American and Caucasian American college students. *Psychiatr Q* 2004; 75:361–73
91. Shumway M, Sentell T, Unick G, Bamberg W: Cognitive complexity of self-administered depression measures. *J Affect Disord* 2004; 83:191–8
92. Moritz S, Meier B, Hand I, Schick M, Jahn H: Dimensional structure of the Hamilton depression rating scale in patients with obsessive-compulsive disorder. *Psychiatry Res* 2004; 125:171–80
93. Zigmond AS, Snaith RP: The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983; 67:361–70

## Hamilton Depression Rating Scale (HAMD-17)

**Instructions:** To rate the severity of depression in patients who are already diagnosed as depressed, administer this questionnaire. The higher the score, the more severe the depression.

For each item, circle the number next to the correct item (only one response per item).

### 1. Depressed Mood (sadness, hopeless, helpless, worthless)

- 0 - Absent
- 1 - These feeling states indicated only on questioning
- 2 - These feeling states spontaneously reported verbally
- 3 - Communicates feeling states non-verbally – i.e., through facial expression, posture, voice, and tendency to weep
- 4 - Patient reports VIRTUALLY ONLY these feeling states in his spontaneous verbal and non-verbal communication

### 2. Feelings of Guilt

- 0 - Absent.
- 1 - Self reproach, feels he has let people down
- 2 - Ideas of guilt or rumination over past errors or sinful deeds
- 3 - Present illness is a punishment. Delusions of guilt
- 4 - Hears accusatory or denunciatory voices and/or experiences threatening visual hallucinations

### 3. Suicide

- 0 - Absent
- 1 - Feels life is not worth living
- 2 - Wishes he were dead or any thoughts of possible death to self
- 3 - Suicidal ideas or gesture
- 4 - Attempts at suicide (any serious attempt rates 4)

### 4. Insomnia Early

- 0 - No difficulty falling asleep
- 1 - Complains of occasional difficulty falling asleep – i.e., more than 1/2 hour
- 2 - Complains of nightly difficulty falling asleep

### 5. Insomnia Middle

- 0 - No difficulty
- 1 - Patient complains of being restless and disturbed during the night
- 2 - Waking during the night – any getting out of bed rates 2 (except for purposes of voiding)

### 6. Insomnia Late

- 0 - No difficulty
- 1 - Waking in early hours of the morning but goes back to sleep
- 2 - Unable to fall asleep again if he gets out of bed

### 7. Work and Activities

- 0 - No difficulty
- 1 - Thoughts and feelings of incapacity, fatigue or weakness related to activities, work or hobbies
- 2 - Loss of interest in activity, hobbies or work – either directly reported by patient, or indirect in listlessness, indecision and vacillation (feels he has to push self to work or activities)

- 3 - Decrease in actual time spent in activities or decrease in productivity
- 4 - Stopped working because of present illness

**8. Retardation: Psychomotor** (slowness of thought and speech; impaired ability to concentrate; decreased motor activity)

- 0 - Normal speech and thought
- 1 - Slight retardation at interview
- 2 - Obvious retardation at interview
- 3 - Interview difficult
- 4 - Complete stupor

**9. Agitation**

- 0 - None
- 1 - Fidgetiness
- 2 - Playing with hands, hair, etc.
- 3 - Moving about, can't sit still.
- 4 - Hand wringing, nail biting, hair-pulling, biting of lips.

**10. Anxiety (psychological)**

- 0 - No difficulty
- 1 - Subjective tension and irritability
- 2 - Worrying about minor matters
- 3 - Apprehensive attitude apparent in face or speech
- 4 - Fears expressed without questioning

**11. Anxiety Somatic:** Physiological concomitants of anxiety (i.e., effects of autonomic overactivity, "butterflies," indigestion, stomach cramps, belching, diarrhea, palpitations, hyperventilation, paresthesia, sweating, flushing, tremor, headache, urinary frequency). Avoid asking about possible medication side effects (i.e., dry mouth, constipation)

- 0 - Absent
- 1 - Mild
- 2 - Moderate
- 3 - Severe
- 4 - Incapacitating

**12. Somatic Symptoms (gastrointestinal)**

- 0 - None.
- 1 - Loss of appetite but eating without encouragement from others. Food intake about normal
- 2 - Difficulty eating without urging from others. Marked reduction of appetite and food intake.

**13. Somatic Symptoms General**

- 0 - None
- 1 - Heaviness in limbs, back or head. Backaches, headache or muscle aches. Loss of energy and fatigability.
- 2 - Any clear-cut symptom rates "2"

**14. Genital Symptoms** (symptoms such as loss of libido; impaired sexual performance; menstrual disturbances)

- 0 - Absent
- 1 - Mild
- 2 - Severe

**15. Hypochondriasis**

- 0 - Not present
- 1 - Self-absorption (bodily)
- 2 - Preoccupation with health
- 3 - Frequent complaints, requests for help, etc.
- 4 - Hypochondriacal delusions

**16. Loss of Weight**

- 0 - No weight loss
- 1 - Probable weight loss associated with present illness
- 2 - Definite (according to patient) weight loss
- 3 - Not assessed

**17. Insight**

- 0 - Acknowledges being depressed and ill
- 1 - Acknowledges illness but attributes cause to bad food, climate, overwork, virus, need for rest, etc.
- 2 - Denies being ill at all

**Total Score (total of circled responses): \_\_\_\_\_**

## Montgomery Asberg Depression Rating Scale

### 1. Apparent Sadness

Representing despondency, gloom and despair (more than just ordinary transient low spirits) reflected in speech, facial expression, and posture. Rate by depth and inability to brighten up.

- 0 - No sadness.
- 2 - Looks dispirited but does brighten up without difficulty.
- 4 - Appears sad and unhappy most of the time.
- 6 - Looks miserable all the time. Extremely despondent.

### 2. Reported Sadness

Representing reports of depressed mood, regardless of whether it is reflected in appearance or not. Includes low spirits, despondency or the feeling of being beyond help and without hope.

- 0 - Occasional sadness in keeping with the circumstances.
- 2 - Sad or low but brightens up without difficulty.
- 4 - Pervasive feelings of sadness or gloominess. The mood is still influenced by external circumstances.
- 6 - Continuous or unvarying sadness, misery or despondency.

### 3. Inner Tension

Representing feelings of ill-defined discomfort, edginess, inner turmoil, mental tension mounting to either panic, dread or anguish. Rate according to intensity, frequency, duration and the extent of reassurance called for.

- 0 - Placid. Only fleeting inner tension.
- 2 - Occasional feelings of edginess and ill-defined discomfort.
- 4 - Continuous feelings of inner tension or intermittent panic which the patient can only master with some difficulty.
- 6 - Unrelenting dread or anguish. Overwhelming panic.

### 4. Reduced Sleep

Representing the experience of reduced duration or depth of sleep compared to the subject's own normal pattern when well.

- 0 - Sleeps as normal.
- 2 - Slight difficulty dropping off to sleep or slightly reduced, light or fitful sleep.
- 4 - Moderate stiffness and resistance.
- 6 - Sleep reduced or broken by at least 2 hours.

### 5. Reduced Appetite

Representing the feeling of a loss of appetite compared with when well. Rate by loss of desire for food or the need to force oneself to eat.

- 0 - Normal or increased appetite.
- 2 - Slightly reduced appetite.
- 4 - No appetite. Food is tasteless.
- 6 - Needs persuasion to eat at all.

### 6. Concentration Difficulties

Representing difficulties in collecting one's thoughts mounting to an incapacitating lack of concentration.

- 0 - No difficulties in concentrating.
- 2 - Occasional difficulties in collecting one's thoughts.

- 4 - Difficulties in concentrating and sustaining thought which reduced ability to read or hold a conversation.
- 6 - Unable to read or converse without great difficulty.

**7. Lassitude**

Representing difficulty in getting started or slowness in initiating and performing everyday activities.

- 0 - Hardly any difficulty in getting started. No sluggishness.
- 2 - Difficulties in starting activities.
- 4 - Difficulties in starting simple routine activities which are carried out with effort.
- 6 - Complete lassitude. Unable to do anything without help.

**8. Inability to Feel**

Representing the subjective experience of reduced interest in the surroundings or activities that normally give pleasure. The ability to react with adequate emotion to circumstances or people is reduced.

- 0 - Normal interest in the surroundings and in other people.
- 2 - Reduced ability to enjoy usual interests.
- 4 - Loss of interest in the surroundings. Loss of feelings for friends and acquaintances.
- 6 - The experience of being emotionally paralyzed, inability to feel anger, grief or pleasure and a complete or even painful failure to feel for close relatives and friends.

**9. Pessimistic Thoughts**

Representing thoughts of guilt, inferiority, self-reproach, sinfulness, remorse, and ruin.

- 0 - No pessimistic thoughts.
- 2 - Fluctuating ideas of failure, self-reproach or self-depreciation.
- 4 - Persistent self-accusations or definite but still rational ideas of guilt or sin. Increasingly pessimistic about the future.
- 6 - Delusions of ruin, remorse or irredeemable sin. Self-accusations which are absurd and unshakable.

**10. Suicidal Thoughts**

Representing the feeling that life is not worth living, that a natural death would be welcome, suicidal thoughts, and preparations for suicide. Suicide attempts should not in themselves influence the rating.

- 0 - Enjoys life or takes it as it comes.
- 2 - Weary of life. Only fleeting suicidal thoughts.
- 4 - Probably better off dead. Suicidal thoughts are common, and suicide is considered as a possible solution, but without specific plans or intentions.
- 6 - Explicit plans for suicide when there is an opportunity. Active preparations for suicide.

**Total Score (total of circled responses): \_\_\_\_\_**

## QIDS-SR16

**Instructions:** Please circle one response to each item that best describes you for the past 7 days.

**During the Past 7 Days...**

### 1. Falling Asleep

- 0 - I never take longer than 30 min to fall asleep.
- 1 - I take at least 30 min to fall asleep, less than half the time.
- 2 - I take at least 30 min to fall asleep, more than half the time.
- 3 - I take more than 60 min to fall asleep, more than half the time.

### 2. Sleep During the Night

- 0 - I do not wake up at night.
- 1 - I have a restless, light sleep with a few brief awakenings each night.
- 2 - I wake up at least once a night, but I go back to sleep easily.
- 3 - I awaken more than once a night and stay awake for 20 min or more, more than half the time.

### 3. Waking Up Too Early

- 0 - Most of the time, I awaken no more than 30 min before I need to get up.
- 1 - More than half the time, I awaken more than 30 min before I need to get up.
- 2 - I almost always awaken at least 1 hour or so before I need to, but I go back to sleep eventually.
- 3 - I awaken at least 1 hour before I need to, and can't go back to sleep.

### 4. Sleeping Too Much

- 0 - I sleep no longer than 7–8 hours/night, without napping during the day.
- 1 - I sleep no longer than 10 hours in a 24-hour period including naps.
- 2 - I sleep no longer than 12 hours in a 24-hour period including naps.
- 3 - I sleep longer than 12 hours in a 24-hour period including naps.

### 5. Feeling Sad

- 0 - I do not feel sad.
- 1 - I feel sad less than half the time.
- 2 - I feel sad more than half the time.
- 3 - I feel sad nearly all of the time.

**Please Complete Either 6 or 7 (Not Both)**

### 6. Decreased Appetite

- 0 - There is no change in my usual appetite.
- 1 - I eat somewhat less often or lesser amounts of food than usual.
- 2 - I eat much less than usual and only with personal effort.
- 3 - I rarely eat within a 24-hour period, and only with extreme personal effort or when others persuade me to eat.

-Or-

**7. Increased Appetite**

- 0 - There is no change from my usual appetite.
- 1 - I feel a need to eat more frequently than usual.
- 2 - I regularly eat more often and/or greater amounts of food than usual.
- 3 - I feel driven to overeat both at mealtime and between meals.

Please Complete Either 8 or 9 (Not Both)

**8. Decreased Weight (Within the Last 2 Weeks)**

- 0 - I have not had a change in my weight.
- 1 - I feel as if I've had a slight weight loss.
- 2 - I have lost 2 pounds or more.
- 3 - I have lost 5 pounds or more.

-Or-

**9. Increased Weight (Within the Last 2 Weeks)**

- 0 - I have not had a change in my weight.
- 1 - I feel as if I've had a slight weight gain.
- 2 - I have gained 2 pounds or more.
- 3 - I have gained 5 pounds or more.

**10. Concentration/Decision Making**

- 0 - There is no change in my usual capacity to concentrate or make decisions.
- 1 - I occasionally feel indecisive or find that my attention wanders.
- 2 - Most of the time, I struggle to focus my attention or to make decisions.
- 3 - I cannot concentrate well enough to read or cannot make even minor decisions.

**11. View of Myself**

- 0 - I see myself as equally worthwhile and deserving as other people.
- 1 - I am more self-blaming than usual.
- 2 - I largely believe that I cause problems for others.
- 3 - I think almost constantly about major and minor defects in myself.

**12. Thoughts of Death or Suicide**

- 0 - I do not think of suicide or death.
- 1 - I feel that life is empty or wonder if it's worth living.
- 2 - I think of suicide or death several times a week for several minutes.
- 3 - I think of suicide or death several times a day in some detail, or I have made specific plans for suicide or have actually tried to take my life.

**13. General Interest**

- 0 - There is no change from usual in how interested I am in other people or activities.
- 1 - I notice that I am less interested in people or activities.
- 2 - I find I have interest in only one or two of my formerly pursued activities.
- 3 - I have virtually no interest in formerly pursued activities.

**14. Energy Level**

- 0 - There is no change in my usual level of energy.
- 1 - I get tired more easily than usual.

- 2 - I have to make a big effort to start or finish my usual daily activities (for example, shopping, homework, cooking or going to work).
- 3 - I really cannot carry out most of my usual daily activities because I just don't have the energy.

**15. Feeling Slowed Down**

- 0 - I think, speak, and move at my usual rate of speed.
- 1 - I find that my thinking is slowed down or my voice sounds dull or flat.
- 2 - It takes me several seconds to respond to most questions and I'm sure my thinking is slowed.
- 3 - I am often unable to respond to questions without extreme effort.

**16. Feeling Restless**

- 0 - I do not feel restless.
- 1 - I'm often fidgety, wring my hands, or need to shift how I am sitting.
- 2 - I have impulses to move about and am quite restless.
- 3 - At times, I am unable to stay seated and need to pace around.

**Total Score\*:** \_\_\_\_\_

\*Total of circled items including *either* 6 or 7, but not both, and *either* 8 or 9 but not both

## Zung Self-Rating Depression Scale

**Instructions:** Please read each statement and decide how much of the time the statement describes how you have been feeling during the past several days. Make a check mark (✓) in the appropriate column.

	A little of the time	Some of the time	Good part of the time	Most of the time
1. I feel down-hearted and blue				
2. Morning is when I feel the best				
3. I have crying spells or feel like it				
4. I have trouble sleeping at night				
5. I eat as much as I used to				
6. I still enjoy sex				
7. I notice that I am losing weight				
8. I have trouble with constipation				
9. My heart beats faster than usual				
10. I get tired for no reason				
11. My mind is as clear as it used to be				
12. I find it easy to do the things I used to				
13. I am restless and can't keep still				
14. I feel hopeful about the future				
15. I am more irritable than usual				
16. I find it easy to make decisions				
17. I feel that I am useful and needed				
18. My life is pretty full				
19. I feel that others would be better off if I were dead				
20. I still enjoy the things I used to do				

**Total Score\*:** \_\_\_\_\_

\*refer to scoring key

## **Zung Self-Rating Depression Scale Scoring Key**

\*A total score is derived by summing the individual item scores (1–4) and ranges from 20 to 80. The items are scored: 1 = a little of the time, through 4 = most of the time, *except* for items 2, 5, 6, 11, 12, 14, 16, 17, 18, and 20 which are scored inversely (4 = a little of the time)