Published in final edited form as:
J Psychiatr Pract. 2009 March ; 15(2): 87–94. doi:10.1097/01.pra.0000348361.03925.b3.

Screening for Major Depression in Private Practice

Ira H. Bernstein, Ph. D.[1], **Burdette Wendt**[2], **Suhayl J. Nasr, MD**[2,3], and **A. John Rush, M.D.**[1,4]

[1] Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, TX

[2] Nasr Psychiatric Services, Michigan City, Indiana

[3] Department of Psychiatry, Indiana University School of Medicine

[4] Duke-NUS Graduate Medical School Singapore

Abstract

Background—Several studies have contrasted the 16-item self-report version of the Quick Inventory of Depressive Symptomatology (QIDS-SR_{16}) with other depression scales, but none has used patients in single, large, private psychiatric practice. This study compared 175 outpatients on the QIDS-SR_{16}, the 17-item Carroll Depression Rating Scale (CDRS-SR_{17}, a self-report modification of the Hamilton Rating Scale for Depression), and the thirteen depression items from the Symptom Check List-90 (SCL-D_{13}). The Mini version of the Structured Clinical Interview for DSM-IV (MiniSCID) served as a "gold standard" to assess depression.

Methods—Basic Item and scale statistics were obtained using classical test theory. Dimensionalities were obtained using factor analysis. Test information functions obtained from the Samejima item response theory model provided additional reliability-like results. This model was also used to compare patients classified as depressed vs. nondepressed on the basis of the MiniSCID. Additional validity information was assessed comparing: (a) ANOVA effect sizes, (b) receiver operating characteristic curves, (c) univariate logistic regression, (d) the MANOVA, and (e) multivariate logistic regression.

Results—The QIDS-SR_{16} related most strongly to MiniSCID diagnoses. The SCL-D_{13}, however, was the most reliable of the three scales ($\alpha = .91$). It was the most sensitive to differences in depression for all but the most depressed patients, for whom the CDRS-SR_{17} was the most sensitive.

Conclusions—All three measures performed satisfactorily, but there are clearly defined advantages to using the QIDS-SR_{16}, as, by its very design, it assesses the core symptoms of depression and does not require a clinician.

Keywords

depression; symptom measurement; Hamilton Depression Rating Scale; Quick Inventory of Depressive Symptomatology–Self-report; Symptom Check List-80; psychometrics; private clinic sample

Introduction

There is a clearly defined need to screen patients' depressive symptoms in private practice settings. In previous research, the 16 item self-report and clinician versions of the Quick Inventory of Depressive Symptomatology (QIDS-SR_{16} and QIDS-C_{16})[1–3] were compared to

Corresponding author: Ira H. Bernstein, Ph. D. Department of Clinical Sciences University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390-9066, Tel.: 214-648-9543; fax: 214-648-3934, Ira.Bernstein@utsouthwestern.edu.

several other inventories in their ability to screen using various populations primarily created for research purposes. For example, Rush et al.[4] compared these scales to the 17-item Hamilton Rating Scale for Depression[5,6] in the course of the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study.[7,8]

One problem with studies like Rush et al. is that the patients were all part of one or another research protocol so additional assumptions must be made if one wishes to generalize to a typical clinic populations. This report presents results obtained with an outpatient clinic population. In particular, the goal is to compare the QIDS-SR_{16}, the 17 item Carroll Depression Rating Scale[9] (CDRS-SR_{17}), and the thirteen depression items from the Symptom Check List-90[10] (SCL-D_{13}). The CDRS-SR_{17}[9] is a self-report version of the Hamilton Rating Scale for Depression.[5,6] Evidence for the comparability of the two instruments may be found in Nasr et al.[11]

We had available diagnoses in this sample as to whether the patient was in a current major depressive episode (MDE) based upon the MiniSCID,[12] which could be used as a "gold standard" against which the former three scales could be compared.

Method

Evaluable Sample

Patients were obtained through a chart review of 224 individuals who made an initial outpatient screening visit at the Nasr Psychiatric Services clinic between 09/2006 and 11/2007. The study was conducted in accordance with international guidelines for good clinical practice and the Declaration of Helsinki and approved by the Institutional Review Board of St. Anthony Memorial Health Centers in Michigan City, Indiana. All patients provided written, informed consent for inclusion of their data in this study. The final sample consisted of 175 who had completed the MiniSCID, QIDS-SR_{16}, CDRS-SR_{17}, and SCL-D_{13}. Their composition is described below under "Sample Characteristics"..

Assessments

Every patient was asked to complete the MiniSCID[12] at a computer terminal in the office followed by paper-and-pencil completion of the SCL-D_{90}, CDRS-SR_{17}, and QIDS-SR_{16} in the same time frame. At a later date, a face-to-face clinical interview was conducted with a board-certified psychiatrist with extensive training and great experience in structured interviews for depression (coauthor SJN). A DSM IV-R diagnosis was generated at this session for clinical and billing purposes based upon clinical judgment derived from the above data. Patients were given the papers together and allowed to do them in whatever order they preferred. One limitation of the study is that it is not possible to determine whether or not there was an order effect since there is no way to identify the order in which the tests were taken. Another limitation of the study was that the clinician who evaluated the patient had access to the test data, but this, in itself, did not imply that any one was relied on most heavily in the diagnostic process.

Test Scoring

All scoring was done in the conventional manner. In particular, the QIDS-SR_{16} was scored by selecting the maximum (most pathological) response to four sleep items, four weight/appetite items, and two restlessness/agitation items, thus converting the sixteen individual items into four domains with five additional domains defined by a single item each for a total of nine. In contrast, all item scores for the CDRS-SR_{17} and SCL-D_{13} were based upon responses to individual items.

Statistical Methods

Analysis proceeded in the following steps.

Classical test theory (CTT) analysis generated the mean and item/total correlation (r_{it}) for each item or domain, coefficient α, and scale means and standard deviations. Correlations among the three measures were also obtained.

Then, component analysis and parallel analyses[13–16] were used to define the number of dimensions on each scale. In a unidimensional scale, the first real principal component's eigenvalue should exceed the first randomly generated principal component's eigenvalue, but the reverse should be true of all subsequent eigenvalues.

The Samejima[17,18] item response theory (IRT) model for graded measures was applied. This involves estimating the parameters of a series of logistic ogives for each item or domain. The purpose here was to obtain the test information functions of each scale. This function describes the sensitivity to differences in the value of depression as a function of its magnitude. In effect, it is a reliability-like measure.

Three analyses examined the relation of the three measures to diagnose patients in an MDE vs. those not in an MDE based upon the MiniSCID. The first of these was simply the effect magnitudes or difference in mean score for currently depressed and currently not depressed groups divided by the pooled within-group standard deviation. This and the associated ANOVA main effect of group upon each of the scales employs linear criteria in assessing the ability of the scales to discriminate depressed vs. nondepressed groups in this study. They are also univariate evaluations since they ignore the covariances among the three measures. Second, univariate logistic regressions were conducted in which the ability of each scale to classify the two groups was determined separately for each scale. Associated with this was an receiver operating characteristic (ROC) analysis. ROC analyses are loglinear since they utilize the log odds ratios rather than probabilities themselves as criteria. Third, the contribution of each measure to the MANOVA discriminant axis was obtained. This is also linear, but it is multivariate since it assesses the ability of a given scale to increment the other two in classification, thus holding the latter constant. Finally, logistic regressions were conducted in which all three scales and pairs of scales were jointly entered as predictors, which assesses the ability of each scale to relate to the log odds of classification holding the two other scales constant. This is both loglinear and multivariate. The various methods are used because it is quite possible for one measure to be most successful by one criterion and other measures to be most successful by different criteria.

Finally, scores on the three scales were equated using methods previously described.[19–21] This allows users who are only familiar with one or two of the scales to place the unfamiliar one(s) in a familiar context.

Results

Sample Characteristics

The patients all came from a clinic located in a suburban part of the Metropolitan Chicago area. They were generally middle class. Most had private insurance (20% were on Medicare) and had some college education. Patients were accepted only upon referral from a primary care colleague or therapists who were well known to the practice. The database originally contained 224 records, of which 16 were excluded as they were younger than 18 years of age, and an additional 33 were excluded because they failed to answer one or more questions on the QIDS-SR_{16}, CDRS-SR_{17}, or SCL-D_{13}, the remaining 175 were included in the subsequent analyses. Those included obtained lower scores on the QIDS-SR_{16} (9.7 vs. 11.9, $p < .01$), did not differ

on the CDRS-SR_{17}, (14.4 vs. 14.4), and obtained lower scores on the SCL-D_{13} (20.2 vs. 22.6, $p < .05$) from those 18 or over who were excluded. The patients' mean age of those included was 44.1 years with a standard deviation of 14.9 years. Their age range was from 18 to 85 years. They did not differ in age from patients over 18 who were excluded (41.9 vs. 41.6). The sample was 41.7% male and 95% Caucasian. Based upon the MiniSCID, 82 patients were classified as clinically depressed, 75 patients were classified as not depressed, and data were missing on 18 patients.

CTT analysis

Tables 1–3 summarize the classical test theory analyses for the QIDS-SR_{16}, CDRS-SR_{17}, and the SCL-D_{13}, respectively. The SCL-D_{13} was the most reliable ($\alpha = 0.90$), followed by the CDRS-SR_{17} ($\alpha = 0.84$), and the QIDS-SR_{16} ($\alpha = 0.81$). The latter two differ minimally. The rank ordering of these values of α is consistent with the differences in scale standard deviation, as expected.

Total scores on the QIDS-SR_{16} correlated .86 with total scores on the CDRS-SR_{17}. This correlation is effectively 1.0 when corrected for unreliability (disattenuated).[22] Similarly, scores on the QIDS-SR_{16} correlated .68 with SCL-D_{13} scores. Total scores on the CDRS-SR_{17} correlated .63 with SCL-D_{13} scores. These latter two correlations became .74 and .73 when disattenuated.

Scale Dimensionality

Figures 1a, 1b, and 1c respectively contain the scree plots for the QIDS-SR_{16}, CDRS-SR_{17}, and SCL-D_{13}. Using the previously cited parallel analysis criterion for imensionality, the QIDS-SR_{16} and SCL-D_{13} are unidimensional. However, the CDRS-SR_{17} is not.

Test Information Functions

Figure 2 contains the test information functions for the three measures; these functions describe the ability of a scale to detect differences in the magnitude of the trait under study (depression in this case). The abscissa is in z-score units so 0 represents the sample mean, +1 is one standard deviation above this mean, etc. As can be seen, the SCL-D_{13} is the most sensitive up to about +2 z-score units above the mean, which incorporates nearly all cases. This is consistent with the fact that its coefficient α was the largest of the three. The CDRS-SR_{17} was the most sensitive to those most highly depressed. Note, however, that these are measures of reliability and not validity.

Diagnostic Validity Based Upon Effect Sizes and ANOVA

The effect sizes were defined as the mean difference between the depressed and nondepressed groups divided by their pooled standard deviation. The respective values for the QIDS-SR_{16}, CDRS-SR_{17}, and SCL-D_{13} were 1.37, 1.21, and .95 thus suggesting the QIDS-SR_{16} is the most discriminating of depressed vs. not depressed groups. ANOVAs conducted in conjunction with these effect sizes indicated that all three differences were significantly greater than zero, $F(1,166) = 73.69$ for the QIDS-SR_{16}, 57.07 for the CDRS-SR_{17}, and 35.46 for the SCL-D_{13}, all $ps < .0001$.

Diagnostic Validity Based Upon Univariate Logistic Regression and ROC Analysis

A logistic regression with only the intercept entered provided a residual chi-square of 217.34 on 167 df. Individually, the QIDS-SR_{16}, CDRS-SR_{17}, and SCL-D_{13} reduced this residual by 58.98, 48.17, and 31.90. These decreases are significant on 1 df, $ps < .001$. Thus, the QIDS-SR_{16} again was the most discriminating of the three based on the MiniSCID diagnoses. The SCL-D_{13} was the least discriminating.

Figure 3 contains the ROC curve derived from these three analyses. As can be seen, the QIDS-SR_{16} is most sensitive from a false alarm rate of .0 (specificity of 1.0) up (down) to a false alarm rate of .2 (specificity of .8) past which point the three curves converge. Conversely, the SCL-D_{13} is the least sensitive to this point.

Table 4 contains thresholds, sensitivities, and specificities for selected points along each measure's continua. These points are low (the lowest score for which the sensitivity was at least 30%), medium (the lowest score for which the sensitivity was at least 50%), high (the lowest score for which the sensitivity was at least 70%) and very high (the lowest score for which the sensitivity was at least 90%). The areas under these curves (c-statistic) were .814, .800, and .744 for the QIDS-SR_{16}, CDRS-SR_{17}, and SCL-D_{13}, respectively.

Diagnostic Validity Based Upon MANOVA

A MANOVA utilizing all three scales produce a Wilkes' Λ of 25.31 on 3 and 153 df, $p < .001$. Of greater importance is the discriminant axis, whose weights were .013, .002, and .001 for the QIDS-SR_{16}, CDRS-SR_{17}, and SCL-D_{13}, respectively. This is also consistent with the fact that the QIDS-SR_{16} increments prediction the most while controlling for the other two scales, whereas the SCL-D_{13} increments the least.

Diagnostic Validity Based upon Multivariate Logistic Regression

When the three scales were all combined in a series of logistic regression equations, the QIDS-SR_{16} significantly incremented the joint contribution of the CDRS-SR_{17} and the SCL-D_{13}, $\chi^2(1) = 8.41$, $p < .01$. However, the CDRS-SR_{17} did not increment the joint contribution of the QIDS-SR_{16} and SCL-D_{13}, and the SCL-D_{13} did not increment the joint contribution of the QIDS-SR_{16} and CDRS-SR_{17}, $\chi^2(1) = 1.03$ and < 1, ns. The QIDS-SR_{16} incremented the CDRS-SR_{17} and SCL-D_{13} scales individually, $\chi^2(1) = 12.05$ and 27.97, ps $< .001$, the CDRS-SR_{17} incremented the SCL-D_{13}, $\chi^2(1) = 20.59$, $p < .001$, and the SCL-D_{13} incremented the CDRS-SR_{17}. $\chi^2(1) = 4.33$, $p < .05$.

Equated Scale Scores

Table 5 shows how the scores for the three scales may be equated, using thresholds of non, mild, moderate, severe, and very severe defined using the QIDS-SR_{16} total score ($\leq$5, 6–10, 11–15, 16–20, 21+).

Discussion

The main finding is that although the SCL-D_{13} was the most reliable in both the CTT and IRT senses, it was the least valid. Conversely, the QIDS-SR_{16} was the most valid based on four different analyses (effect size/ANOVA, univariate logistic regression/ROC analysis, MANOVA, and multivariate logistic regression) - albeit only slightly more so. The multidimensionality of the CDRS-SR_{17} might be an advantage as it is potentially sensitive to both the signs and symptoms of depression, thus affording it two opportunities to detect depression. In contrast, the unidimensionality of the QIDS-SR_{16} suggests that it is limited to the symptoms of depression. It is not surprising that the CDRS-SR_{17} was multidimensional, as we have previously observed that the Hamilton Rating Scale for Depression upon which it is based, is also multidimensional.[4,19] However, the two scales basically correlate perfectly within the limits imposed by their unreliability. In contrast, both are somewhat different from the SCL-D_{13}. As one would expect from the fact that there is substantial similarity among the three sets of items, these differences tend to be small.

The QIDS-SR_{16} and CDRS-SR_{17} item means are also fairly consistent with values we have reported for other samples.[4,19] The QIDS-SR_{16} is a very user friendly tool for busy

practitioners. It can be scored in less than a minute and can give a rapid view of the core symptom domains of depression. It has also clearly been shown to be unidimensional in various settings (e.g., Bernstein *et al*.,[23] Carmody *et al*..[20] Rush *et al*.[4]). By its very design, its items clearly reflect the core DSM criteria for depression. The CDRS-SR_{17} is a self-rated version used in this practice and only takes about 2 minutes to score by an experienced assistant. However, It does not relate as clearly to these core symptom domains. The SCL-D_{13} takes longer to complete for both the patient and the scorer. The scores can be plotted on a graph which can visually indicate the areas of elevation. It too does not relate clearly to the core symptom domains.

Limitations in this report are noteworthy. The diagnostic evaluation was not completely blind to the test results. The population is small and representative of a single practice. On the other hand, diagnoses were based on a structured interview, and the patients were typical of psychiatric practices. Finally the order of test administration was not randomized.

In summary, given the slightly greater validity of the QIDS-SR_{16} compared to the CDRS-SR_{17} and the SCL-D_{13} by all criteria, it should become the tool of choice in private practice settings although all three measures are acceptable.

Acknowledgments

This project was funded in part by the National Institute of Mental Health (NIMH) grants MH-68852 to the University of Texas at Arlington, Ira H. Bernstein, Ph.D., PI and MH-68851 to the University of Texas Southwestern Medical Center at Dallas, A. John Rush, M.D., PI.

References

1. Rush AJ, Carmody TJ, Reimitz PE. The Inventory of Depressive Symptomatology (IDS): clinician (IDS-C) and self-report (IDS-SR) ratings of depressive symptoms. Int J Methods Psychiatr Res 2000;9:45–59.
2. Rush AJ, Trivedi MH, Ibrahim HM, et al. 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. [PubMed: 12946886]Erratum p. 585
3. Trivedi MH, Rush AJ, Ibrahim HM, et al. 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. [PubMed: 14971628]
4. Rush AJ, Bernstein IH, Trivedi MH, et al. An evaluation of the Quick Inventory of Depressive Symptomatology and the Hamilton Rating Scale for Depression: a Sequenced Treatment Alternatives to Relieve Depression trial report. Biol Psychiatry 2006;59:493–501. [PubMed: 16199008]Epub 2005 Sep 30
5. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62. [PubMed: 14399272]
6. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278–96. [PubMed: 6080235]
7. Fava M, Rush AJ, Trivedi MH, et al. Background and rationale for the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study. Psychiatr Clin North Am 2003;26:457–94. x. [PubMed: 12778843]
8. Rush AJ, Fava M, Wisniewski SR, et al. Sequenced Treatment Alternatives to Relieve Depression (STAR*D): rationale and design. Control Clin Trials 2004;25:119–42. [PubMed: 15061154]
9. Carroll BJ, Feinberg M, Smouse PE, et al. The Carroll rating scale for depression. I. Development, reliability and validation. Br J Psychiatry 1981;138:194–200. [PubMed: 7272609]
10. Derogatis LR, Rickels K, Rock AF. The SCL-90 and the MMPI: a step in the validation of a new self-report scale. Br J Psychiatry 1976;128:280–9. [PubMed: 1252693]

11. Nasr SJ, Altman EG, Rodin MB, et al. Correlation of the Hamilton and Carroll Depression Rating Scales: a replication study among psychiatric outpatients. J Clin Psychiatry 1984;45:167–8. [PubMed: 6715289]
12. Spitzer, RL.; Williams, JBW.; Gibbon, M., et al. Structured Clinical Interview for DSM-III-R: The Mini-SCID. Washington, DC: American Psychiatric Press; 1992.
13. Horn JL. An empirical comparison of various methods for estimating common factor scores. Educ Psychol Meas 1965;25:313–22.
14. Humphreys LG, Ilgen D. Note on a criterion for the number of common factors. Educ Psychol Meas 1969;29:571–8.
15. Humphreys LG, Montanelli RG Jr. An investigation of the parallel analysis criterion for determining the number of common factors. Multivariate Behav Res 1975;10:193–206.
16. Montanelli RG Jr, Humphreys LG. Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: a Monte Carlo study. Psychometrika 1976;41:341–8.
17. Samejima F. Estimation of latent ability using a response pattern of graded scores. Psychol Monogr 1969;4:2.
18. Samejima, F. Graded response model. In: van Linden, W.; Hambleton, RK., editors. Handbook of Modern Item Response Theory. New York: Springer-Verlag; 1997. p. 85-100.
19. Carmody TJ, Rush AJ, Bernstein IH, et al. The Montgomery Asberg and the Hamilton ratings of depression: A comparison of measures. Eur Neuropsychopharmacol 2006;16:601–11. [PubMed: 16769204]
20. Carmody TJ, Rush AJ, Bernstein IH, et al. Making clinicians lives easier: Guidance on use of the QIDS self-report in place of the MADRS. J Affect Disord 2006;95:115–8. [PubMed: 16707162]
21. Orlando M, Sherbourne CD, Thissen D. Summed-score linking using item response theory: application to depression measurement. Psychol Assess 2000;12:354–9. [PubMed: 11021160]
22. Nunnally, JC.; Bernstein, IH. Psychometric Theory. Vol. 3. New York: McGraw-Hill; 1994.
23. Bernstein IH, Rush AJ, Stegman D, et al. A comparison of the QIDS-$C_{16,}$ QIDS-SR_{16}, and $MADRS_{10}$ in an adult outpatient clinical sample. Int J Neuropsychopharmacol. 2008submitted (IntJNP-08-0168)

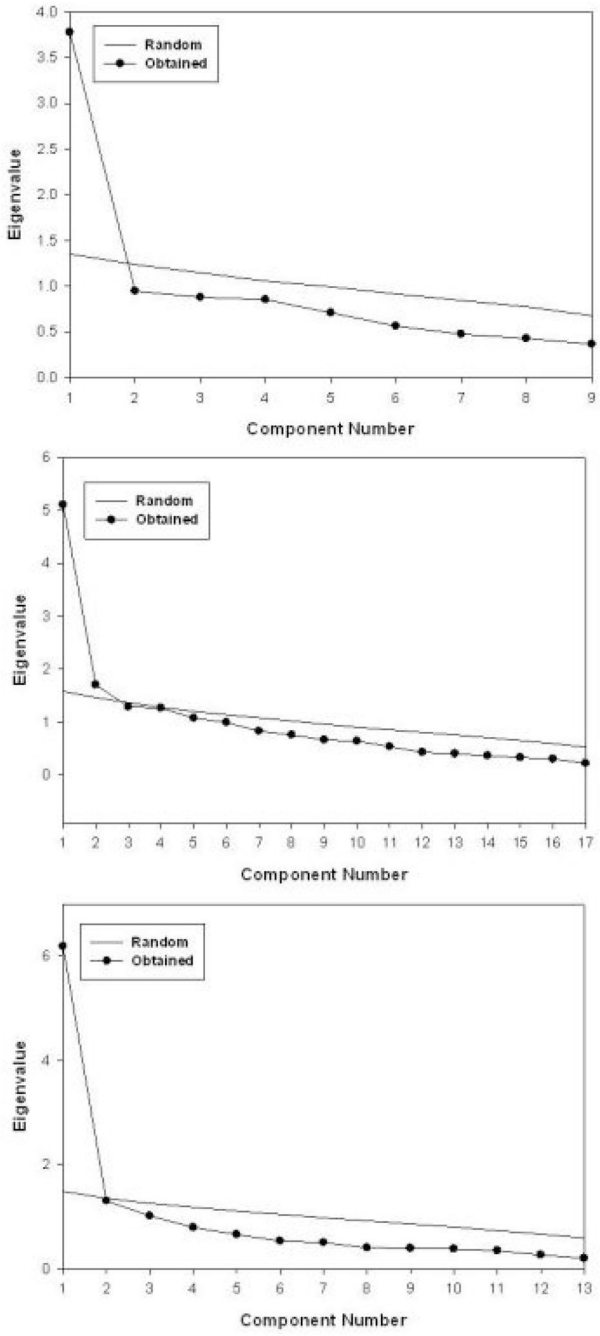

Figure 1.
Scree plots of the QIDS-SR_{16} (top panel), the CDRS-SR_{17} (middle panel), and the SCL-R_{13} (bottom panel)

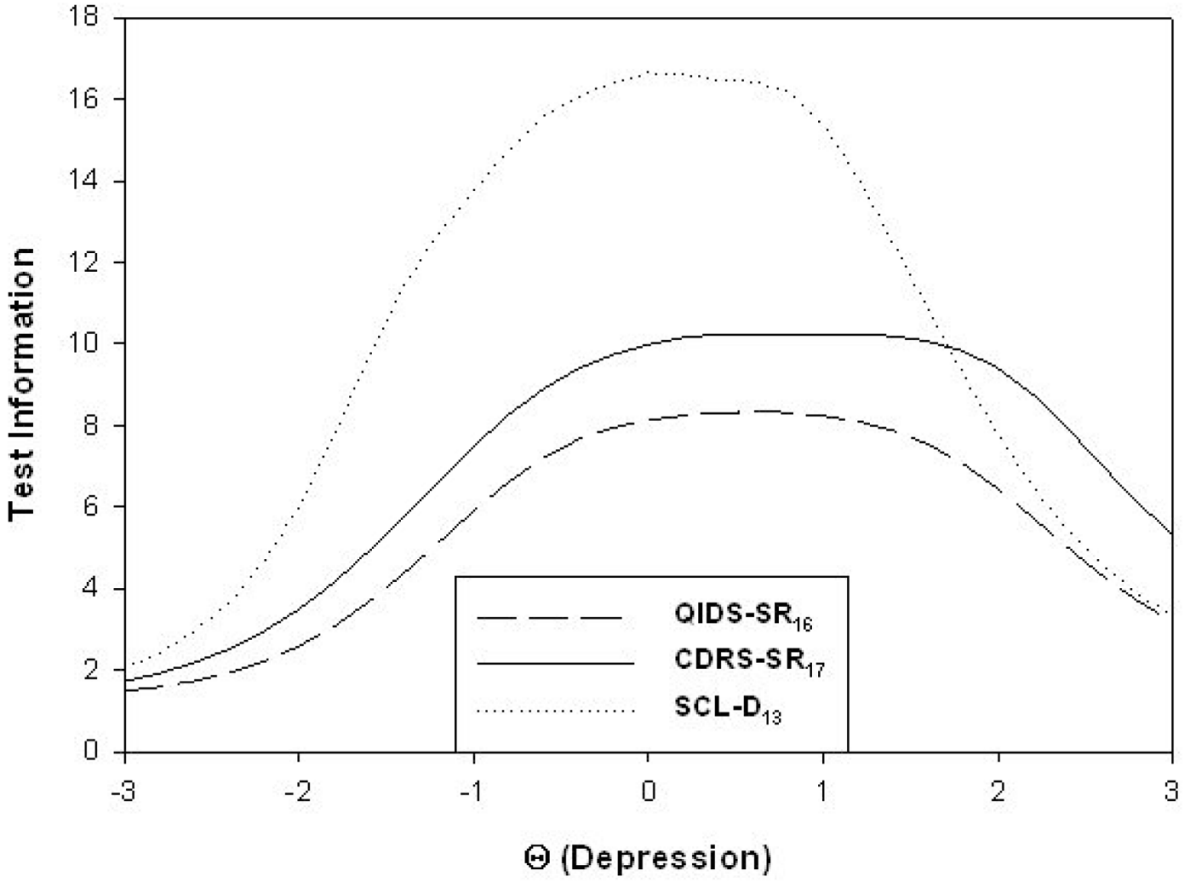

Figure 2.
Test information functions

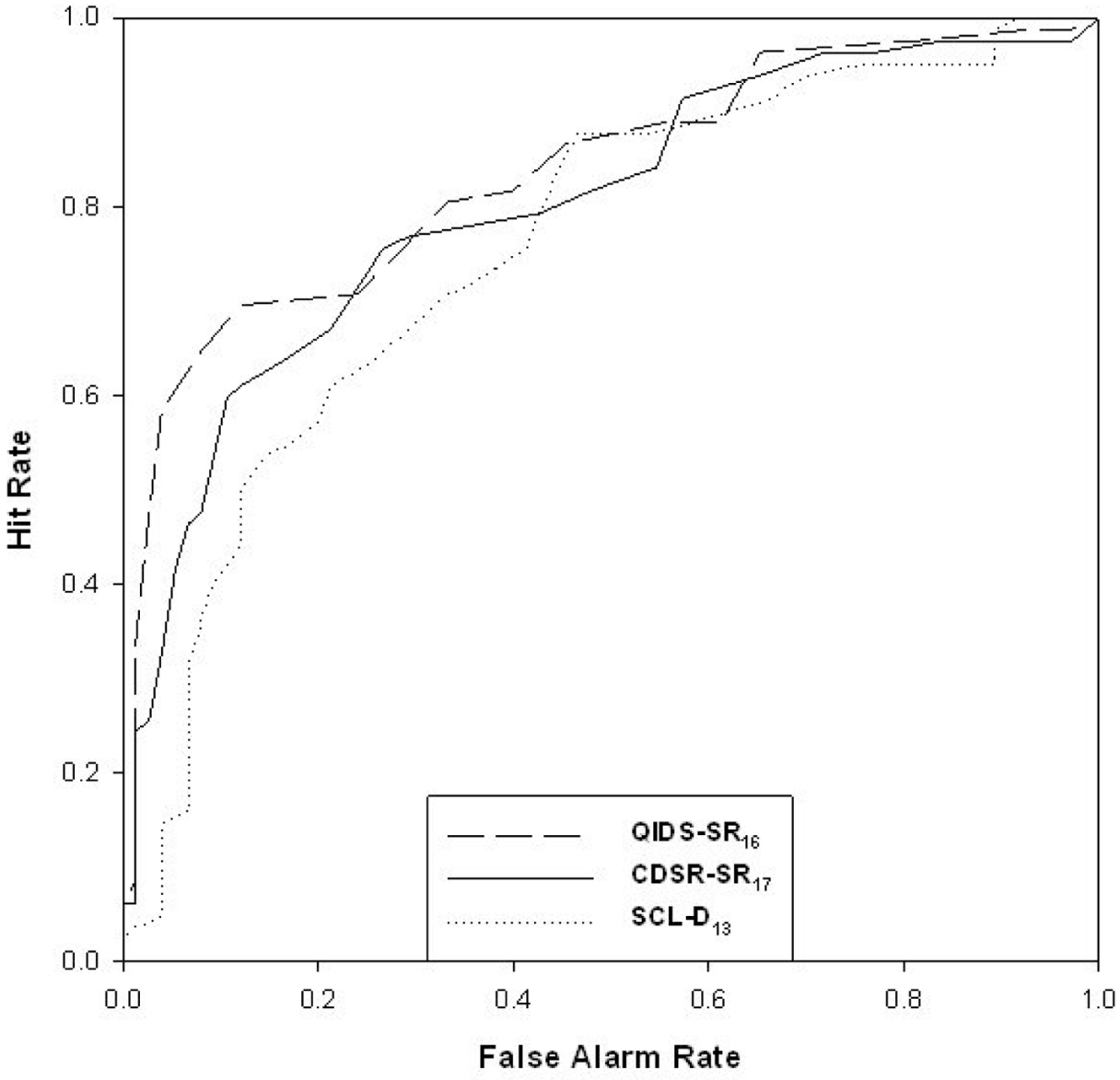

Figure 3.
ROC curves

Table 1

Domain means, item/total correlations, scale internal consistency (coefficient α), scale mean, and scale standard deviation for the QIDS-SR_{16}

Domain	Mean	r_{it}
1. Sleep	2.19	.33
2. Sad Mood	1.42	.65
3. Appetite	1.37	.23
4. Concentration/Decision Making	1.14	.59
5. Self View	.89	.52
6. Thoughts of Death or Suicide	.42	.42
7. General Interest	1.07	.59
8. Energy Level	1.10	.63
9. Restlessness/Agitation	.97	.65
α	.81	
Scale Mean	10.57	
Scale SD	5.54	

Table 2

Domain means, item/total correlations, scale internal consistency (coefficient α), scale mean, and scale standard deviation for the CDRS-SR_{17}

Item	Mean	r_{it}
1. Depressed mood	1.21	.69
2. Guilt Mood	1.11	.48
3. Suicidal feelings	.35	.41
4. Initial insomnia	.91	.37
5. Middle insomnia	1.12	.37
6. Delayed insomnia	.79	.24
7. Work and interests	1.50	.66
8. Retardation	1.42	.63
9. Agitation	1.18	.54
10. Psychic anxiety	1.41	.67
11. Somatic anxiety	.85	.40
12. GI symptoms	.42	.51
13. Somatic symptoms	1.18	.51
14. Libido	.67	.38
15. Hypochondriasis	.82	.41
16. Loss of insight	.54	.11
17. Loss of weight	.22	.00
α	.84	
Scale Mean	15.70	
Scale SD	8.69	

Table 3

Domain means, item/total correlations, scale internal consistency (coefficient α), scale mean, and scale standard deviation for the SCL-D_{13}

Item	Mean	r_{it}
1. Loss of sexual interest or pleasure	1.66	.37
2. Feeling low in energy or slowed down	2.22	.65
3. Thought of ending your life	.68	.58
4. Crying easily	1.46	.45
5. Feelings of being trapped or caught	.93	.52
6. Blaming yourself for things	1.85	.61
7. Feeling lonely	1.91	.69
8. Feeling blue	2.31	.79
9. Worrying too much about things	2.57	.60
10. Feeling no interest in things	1.83	.65
11. Feeling hopeless about the future	1.61	.69
12. Feeling everything is an effort	1.51	.68
13. Feelings of worthlessness	1.52	.73
α	.90	
Scale Mean	22.05	
Scale SD	12.08	

Table 4

Threshold scores (Thresh.), Sensitivities (Sens.), and Specificities (Spec.) at Four Levels of Severity for the QIDS-SR_{16}, CDRS-SR_{17}, and SCL-D_{13}

	QIDS-SR_{16}			CDRS-SR_{17}			SCL-D_{13}		
Level	**Thresh.**	**Sens.**	**Spec.**	**Thresh.**	**Sens.**	**Spec.**	**Thresh.**	**Sens.**	**Spec.**
Low	11	30	13	15	32	21	20	30	34
Medium	14	52	3	20	52	8	27	50	11
High	16	73	1	24	74	3	34	73	6
Very High	19	90	1	31	92	1	40	90	4

Note: Low is the first level of each test for which the sensitivity was ≥30%; Medium is the first level of each test for which the sensitivity was ≥50%, High is the first level of each test for which the sensitivity was ≥70%, and Very High is the first level of each test for which the sensitivity was 90% or higher.

Table 5

Equated Scale Scores on the QIDS-SR_{16}, CDRS-SR_{17}, and SCL-D_{13}

Θ	QIDS-SR_{16}	CDRS-SR_{17}	SCL-D_{13}
−2.0	0		1
−1.3	3	3	6
−.9	5	6	10
.0	10	15	22
.4	13	19–20	28
.7	15	24	32–33
1.0	17	28	37
1.1	18	29–30	38
1.4	20	33–34	42
1.9	23	39	47
2.8	27	47	52