

# The Quick Inventory of Depressive Symptomatology, Adolescent Version (QIDS-A<sub>17</sub>): A Psychometric Evaluation

Charlotte L Haley<sup>1</sup>, Betsy D Kennard<sup>1,2</sup>, David W Morris<sup>1</sup>, Ira H Bernstein<sup>3</sup>, Thomas Carmody<sup>4</sup>, Graham J Emslie<sup>1,2</sup>, Taryn L Mayes<sup>1</sup>, A John Rush<sup>5-7</sup>

<sup>1</sup>Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX, USA; <sup>2</sup>Children's Health, Children's Medical Center of Dallas, Dallas, TX, USA; <sup>3</sup>Department of Psychology, The University of Texas at Arlington, Arlington, TX, USA; <sup>4</sup>Peter O'Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, Dallas, TX, USA; <sup>5</sup>Department of Psychiatry and Clinical Sciences, Duke-National University of Singapore, Singapore, Singapore; <sup>6</sup>Department of Psychiatry and Behavioral Sciences, Texas Tech University – Health Sciences Center, Permian Basin, TX, USA; <sup>7</sup>Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, USA

Correspondence: Betsy D Kennard, University of Texas Southwestern Medical Center, Child Psychiatry, 5323 Harry Hines Blvd, Dallas, TX, 75390-8589, USA, Tel +1 214.645.8680, Fax +1 214.648.3914, Email beth.kennard@utsouthwestern.edu

**Objective:** The current study aimed to evaluate the psychometric features of the Quick Inventory of Depressive Symptomatology, Adolescent version (QIDS-A<sub>17</sub>) and the clinician-rated Children's Depression Rating Scale-Revised (CDRS-R).

**Methods:** Altogether, 103 outpatients (8 to 17 years) completed the self-report QIDS-A<sub>17</sub>-SR. Clinician interviews of adolescents (QIDS-A<sub>17</sub>-C (Adolescent)) and of parents (QIDS-A<sub>17</sub>-C (Parent)) were combined to create the QIDS-A<sub>17</sub>-C(Composite) and the CDRS-R.

**Results:** All QIDS-A<sub>17</sub> measures and the CDRS-R evidenced high total score correlations and internal consistency. Factor analysis found all four measures to be unidimensional. Item Response Theory (IRT) analysis found results that complemented the reliability results found in CTT. All four also demonstrated discriminant diagnostic validity based on logistic regression and ANOVA analyses.

**Conclusion:** The psychometric properties of the self-report and composite versions of the QIDS-A<sub>17</sub> suggest acceptability as a measure of depression in adolescents either as a measure of depressive symptoms or severity of illness in adolescents. The self-report version may be a helpful tool in busy clinical practices.

**Keywords:** pediatric depression, rating scale, psychometric properties, self-report measures

## Introduction

Over 15% of adolescents aged 12 to 17 years in 2018–2019 had a major depressive episode, according to data from the Youth Risk Behavior Surveillance (YBRS) system,<sup>1</sup> many of which are recurrent<sup>2</sup> with substantial functional impairment and mortality.<sup>3–5</sup> The National Survey of Children's Health found a 27% increase in depression from 2016–2020.<sup>6</sup> The prevalence rates of depression in adolescents increased from 8.3% to 12.9% between 2011 and 2016.<sup>7</sup> Under-recognition and under treatment both contribute to suboptimal outcomes.<sup>8–10</sup>

Measurement-based care (MBC) is superior to clinical judgment alone regarding treatment impact and has been shown to improve treatment outcomes in youth with depression.<sup>10–12</sup> MBC entails the regular clinical use of reliable, valid symptom severity measures that are sensitive to symptom change<sup>13</sup> to assist clinicians in monitoring outcomes and making timely revisions in the treatment regimen when needed.<sup>14,15</sup>

The Children's Depression Rating Scale-Revised (CDRS-R),<sup>16–18</sup> the present standard for measuring depressive symptoms in clinical trials for children and adolescents,<sup>19</sup> has good psychometric properties,<sup>20,21</sup> but there are some limitations.<sup>22</sup> For example, it has poor inter-informant reliability,<sup>23</sup> lacks precision in how to weight data from various informants to create the consensus score, requires training time for raters, and can be costly. The total CDRS-R score

includes ratings of both symptoms and function, which can create a potential confound in measuring symptom remission.<sup>24</sup> Thus, the CDRS-R may be less sensitive to detecting symptom remission since the rated functional difficulties may have been present prior to MDE onset or, as in adults, they may take longer to resolve than depressive symptoms.<sup>25</sup> Further, the CDRS-R lacks a self-report version, which may be especially useful in adolescents since important symptoms of depression (an internalizing disorder) may not come to parental attention until obvious behavioral problems occur.<sup>26,27</sup> Adult studies have found self-report measures can replace more time-consuming clinician-rated measures, with good correlation<sup>2</sup> and concurrent validity<sup>25</sup> established between the adult self-report and clinician-rated measures.<sup>28–30</sup>

The Quick Inventory of Depressive Symptoms (QIDS) is a 16-item measure (self-report and clinician-rated versions) for adults with depression with solid psychometric properties and substantial data supporting sensitivity to change.<sup>31</sup> The QIDS assesses the 9 criterion domains used to diagnose a major depressive disorder.<sup>31,32</sup> The QIDS has been widely used in adults.

The QIDS has not been widely studied in children and adolescents. Using the adult version of the QIDS-SR with minor modifications to take irritability into account as an alternative to sadness to qualify for the diagnosis, Bernstein, Rush, Trivedi et al<sup>33</sup> found this adult-adapted version to be unifactorial with satisfactory psychometric properties using the adult version of the QIDS-SR in a sample of adolescents. More recently, Zhang et al<sup>34</sup> reported satisfactory psychometric performance of the adult QIDS-SR in 246 outpatient youths in China.

While the 9 criterion symptom domains are the same in children and adults based on the DSM-5-TR,<sup>32</sup> the wording of the items in the adult QIDS might not be optimal for either self-report or clinician-rated QIDS for children and adolescents. To address this potential shortcoming, we developed the 17-item version of the QIDS for adolescents (QIDS-A<sub>17</sub>). This report is the first assessment of this instrument.

The present study evaluates the psychometric properties of each QIDS-A<sub>17</sub> (self-report and clinician-rated) as well as the composite versions of the QIDS-A<sub>17</sub> and the CDRS-R (each based on both the parent and adolescent clinical interviews) using classical test theory (CTT) and item response theory analyses.

## Methods

### QIDS-A Scale Development

Based on consultation with child experts prior to developing the scale, several well-established and commonly used assessments to measure emotional symptoms, including depression in children and adolescents, were reviewed. The purpose of the review was to identify valid, reliable, methods of describing symptoms constructs of depression as they manifest in, and are understood by, children and adolescents. The Child Behavior Checklist (CBCL),<sup>35</sup> was included because it was being widely used in the public school system. The Child Depression Inventory (CDI),<sup>36</sup> Child Depression Rating Scale-Revised (CDRS-R),<sup>37</sup> and Patient Health Questionnaire-Adolescent (PHQ-A)<sup>38</sup> were also included as they were modeled after well-established adult depression scales: the Beck Depression Inventory (BDI),<sup>39</sup> Hamilton Depression Rating Scale (HAM-D),<sup>40</sup> and the Patient Health Questionnaire-9 (PHQ-9),<sup>41</sup> respectively.

The original QIDS includes a patient self-report and a clinician rating scale.<sup>31</sup> These forms are appropriate for adults but working with children and adolescents require a different approach. Specifically, both the child and parent need to independently rate the child's level of depressive severity. To accomplish this, six versions of the QIDS-A were created in 2006. The three original QIDS, (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR),<sup>31</sup> were modified to work with children and adolescents, and an additional set of three parallel forms were created for parent/caregiver report.

Questions and item anchors were rewritten to contain age-appropriate language, content, and reading level. Modifications were made to the Likert scale anchors to include additional descriptive language used in the DSM and traditional assessments of child psychopathology. For example, the Mood item anchors were expanded to include feeling down, unhappy, or miserable. Minor changes in person and pronoun were also made throughout (eg, your child, his/hers) for parent versions.

An irritability item was added to maintain consistency with the DSM operational definition of unipolar depression (irritability can be a proxy for depression in children). The scoring was modified to include the Irritability item, again to

maintain consistency with the DSM and the original QIDS; only the highest of the two mood scores (irritability and depression) is used to sum a total score. In this manner, consistent with the original QIDS, only one score for each of the 9 criteria for a MDE is used to calculate a total score, keeping the range of each item (0–3), each domain (0–3) and total scores (0–27) the same as found in the adult QIDS.

## Study Participants

This study recruited children and adolescents from the Division of Child Psychiatry at UT Southwestern and the Child and Adolescent Psychiatry Outpatient Clinic at Children’s Medical Center of Dallas. Study data also included a limited data set from a NIMH-funded study: Pediatric MDD: Sequential Treatment with Fluoxetine and Relapse Prevention CBT (1R01MH-39188; Emslie & Kennard, principal investigators, 2008).

Participants included outpatients 8 to 17 years of age who were still attending school, with no restrictions regarding non-study medications or other treatment(s). Participants with any concurrent general medical condition(s) or Axis I disorder(s) based on clinician diagnosis were included, except as noted below.

Patients were excluded if they had concurrent intellectual disability, active psychosis, terminal illness, neurological disorders that precluded participation from completing study questionnaires, or were unable to speak and read English (the primary self-report and parent-report scales do not have norms for non-English translations). Appropriate approval was obtained from the UT Southwestern Medical Center Institutional Review Board, and all ethical guidelines consistent with the Declaration of Helsinki were followed in the conduct of this research. Participants and caregiver(s) gave informed consent prior to data collection.

## Procedures

Clinicians (doctoral level student, and 4 child psychiatry fellows) interviewed the patient and caregiver separately. The caregiver completed a demographic questionnaire. The patient was first asked to independently complete the QIDS-A<sub>17</sub>-SR. If the patient had difficulty due to age or reading ability, the clinician or research assistant assisted the patient by reading the items verbatim for the patient to answer, giving no additional assistance or guidance. The clinician then interviewed the patient to complete the QIDS-A<sub>17</sub>-C(Adolescent) and the CDRS-R. Next, the clinician interviewed the caregiver to complete the caregiver portion of the clinician version of the QIDS-A<sub>17</sub>-C(Parent) and the CDRS-R. Composite scores for the QIDS-A<sub>17</sub>-C(Composite) and CDRS-R were generated using the clinician’s best estimate of the most valid response based on the ratings of each item (domain) from the parent and the adolescent.

The DSM-IV Checklist<sup>42</sup> (described below) for major depressive episode (MDE) was completed by the clinician to establish the probable presence/absence of an MDE. Participants were assessed only once.

## Measures

### DSM-IV Checklist for MDE

The MDE checklist includes the 9 depression criterion symptoms in the DSM-IV,<sup>42</sup> along with an item assessing level of dysfunction. It is not a structured diagnostic measure, but an indicator of clinical diagnosis based on a clinical interview. The clinician indicates the symptoms for which the patient meets criteria, and then indicates whether the patient meets criteria for a current MDE (*Definite*), is likely depressed but does not meet full criteria for an MDE (*Probable*) or does not meet criteria for an MDE (*No*).

### CDRS-R

The CDRS-R<sup>18</sup> is a 17-item clinician-rated instrument for measuring the presence and severity of depressive symptomatology in children and adolescents over the last 7 days. It is administered to the child and parent separately, and the clinician uses clinical judgment to synthesize the separate responses into a composite score. Each item is rated on a 1 to 5 or 1-to-7-point scale with higher scores indicating more severe symptomatology (The total raw score ranges from 17 to 113). A total score of >40 is indicative of depression, and a score of less than or equal to 28 is frequently used to determine remission.<sup>37</sup>

## QIDS-A<sub>17</sub>

The QIDS-A<sub>17</sub> measures the presence and severity of depressive symptomatology over the last 7 days. In this study, youth completed the self-report (QIDS-A<sub>17</sub>-SR). Clinicians then interviewed the youth and parent separately and rated the QIDS-A<sub>17</sub>-C(Adolescent) and QIDS-A<sub>17</sub>-C(Parent), respectively, then used clinical judgment to synthesize information from the interviews to create the QIDS-A<sub>17</sub>-C(Composite).

The scoring of all QIDS-A<sub>17</sub> versions is identical to the adult QIDS<sub>16</sub> (range 0 to 27) and the items (17 instead of 16: irritability item added to further assess the mood domain) cover the same 9 DSM-IV-TR<sup>32</sup> symptom domains. For domains requiring more than one item (ie, appetite/weight change, sleep disturbance, psychomotor agitation/retardation), the response to the highest scored item in the domain is included in the total score.<sup>31</sup>

## Data Analysis

CTT measures of scale consistency—including Cronbach's alpha,<sup>43</sup> item (symptom) means, total scale score means, item score-total score correlations, and intercorrelations among measures—were computed for all four depression rating scales (QIDS-A<sub>17</sub>-C(Clinician-rated on Adolescent), QIDS-A<sub>17</sub>-C (Composite, clinician-completed based on interviews with parent and child), QIDS-A<sub>17</sub>-SR (adolescent self-report), and CDRS-R (clinician-rated based on both patient and parent interview).

Item response theory (IRT)<sup>44</sup> was used to compare each rating scale and version with respect to item characteristic curves (intercepts and slopes) for each item. These are the IRT equivalents of the item means, and item-total correlations described above in CTT. Uni-dimensionality of the measures was established by parallel analysis.<sup>45</sup> In addition, test information functions (TIF) were computed for each version of each rating to determine and compare which areas of the test are most sensitive.

Univariate logistic regression and univariate analysis of variance (ANOVA) tested diagnostic validity, the ability of each of the four measures to discriminate between depressed and non-depressed participants as defined by the MDE checklist. For purposes of analysis, participants classified as *probably* or *definitely* in an MDE were pooled so that two distinct groups (depressed and non-depressed) were established. Each measure was examined independently for the univariate logistic regression and ANOVA. All results were standardized to control for differences in measurement among the scales. ROC curves<sup>46</sup> were computed for all measures to indicate the sensitivity and specificity of the measures in differentiating depressed from non-depressed participants. The ROC curves<sup>47</sup> were created by plotting the successive values of sensitivity and specificity for the scores on each of the three QIDS-A measures and the CDRS-R. Pairwise comparisons of ROC curves were made using DeLong's Test<sup>48</sup> with a Bonferroni corrected p-value of 0.05/6 = 0.0083. Cronbach's alpha were computed using SPSS. IRT analyses were conducted using MPLUS and the unidimensionality assumption was verified using SAS software. Logistic regression, ANOVA, and ROC analyses were conducted using SAS software.

## Results

### Sample Characteristics

Of the 103 participants, 68 were recruited during a scheduled outpatient clinic visit (no requirements on diagnosis/type of visit), and 35 were included from the NIMH-funded treatment study. Seventy percent of the sample was Caucasian, 51% were female, the mean age was 13.8±2.4 years, and 21 (20%) participants were from 8 to 11 years old. Based on the MDE checklist, 40 (39%) participants met criteria for a current MDE, 55 (53%) were not depressed, and 8 (8%) had some depressive symptoms but did not meet full criteria for a current MDE (Table 1).

## CTT Analyses

### Internal Consistency

Using Cronbach's alpha, medium to high internal consistency was found for all measures. Internal consistency was 0.84 (95% CI: 0.78 to 0.88) for the QIDS-A<sub>17</sub>-C(Adolescent) and 0.84 (95% CI: 0.79 to 0.88) for the QIDS-A<sub>17</sub>-C

**Table 1** Demographic and Clinical Characteristics of Study Participants

Characteristic	All Participants (N = 103)	Participants with an MDE (n=48)	Participant Without an MDE (n=55)	P-value
	Mean ± SD	Mean ± SD	Mean ± SD	
Mean age	13.83 ± 2.4	13.48 ± 2.6	14.14 ± 2.2	0.163
Age group	<b>N (%)</b>	<b>N (%)</b>	<b>N (%)</b>	
Child (≤11 years)	21 (20.4)	14 (29.2)	7 (12.7)	0.039
Adolescent (≥12 years)	82 (79.6)	34 (70.8)	48 (87.3)	
Sex Assigned at Birth				
Male	51 (49.5)	25 (52.1)	26 (47.3)	0.626
Female	52 (50.5)	23 (47.9)	29 (52.7)	
Race/Ethnicity				
Caucasian	73 (70.9)	33 (68.8)	40 (72.7)	0.747
African American	16 (15.5)	8 (16.7)	8 (14.6)	
Asian	1 (1.0)	1 (2.1)	0 (0.0)	
Native American	3 (2.9)	2 (4.2)	1 (1.8)	
Hispanic	10 (9.7)	4 (8.3)	6 (10.9)	
Annual Family Income (n = 98)				
Under \$15,000	9 (9.2)	4 (8.7)	5 (9.6)	0.468
\$15,000-\$35,000	20 (20.4)	10 (21.7)	10 (19.2)	
\$35,000-\$75,000	34 (34.7)	19 (41.3)	15 (28.8)	
Over \$75,000	35 (35.7)	13 (28.3)	22 (42.3)	
Presence of MDE				
Definite presence	40 (38.8)	40 (83.3)	0 (0.0)	<0.001
Probable presence	8 (7.8)	8 (16.7)	0 (0.0)	
No presence	55 (53.4)	0 (0.0)	55 (100.0)	

**Abbreviations:** MDE, Major Depressive Episode; SD, Standard Deviation.

(Composite). The CDRS-R composite had the highest internal consistency ( $\alpha = 0.92$ , 95% CI: 0.90 to 0.94). The QIDS-A<sub>17</sub>-SR internal consistency was 0.78 (95% CI: 0.71 to 0.84).

The Spearman-Brown prophecy formula<sup>49</sup> was computed on the Cronbach's alphas for all QIDS-A<sub>17</sub> measures to correct for any difference due to test length (QIDS-A<sub>17</sub>: 9 domains, CDRS-R: 17 domains) and enable equal comparison of reliability across measures. Such correction increased the reliabilities of the QIDS-A<sub>17</sub> measures to a coefficient alpha of 0.87 for the QIDS-A<sub>17</sub>-SR and 0.91 for both the QIDS-A<sub>17</sub>-C(Adolescent) and the QIDS-A<sub>17</sub>-C(Composite), indicating a high degree of internal consistency for all measures. Compared to the CDRS-R, the difference in uncorrected reliability for the two QIDS-A<sub>17</sub>-C versions was mostly due to the QIDS-A<sub>17</sub>-C versions being shorter scales.

Cronbach's alpha were also computed on the QIDS-A<sub>17</sub> measures with the irritability item removed to examine the item's contribution. Removing the irritability item increased reliability by 0.01 in all QIDS-A<sub>17</sub> measures.

### Item Means and Item-Total Correlations

Item, total score means, and item-total correlations ( $r_{it}$ ) were computed for all measures (Table 2 and Table 3). The item means measured the tendency of participants to endorse particular symptoms. The item-total correlations (domain-total correlations) measured how robustly a symptom relates to overall depression as indicated by the total scale score.

**Table 2** CTT Analysis of the QIDS-A<sub>17-C</sub>(Adolescent), QIDS-A<sub>17-C</sub>(Composite), and QIDS-A<sub>17-SR</sub>

Domain	QIDS-A <sub>17-C</sub> (Adolescent)		QIDS-A <sub>17-C</sub> (Composite)		QIDS-A <sub>17-SR</sub>	
	(n = 101)		(n = 102)		(n = 102)	
	Mean	$r_{it}$	Mean	$r_{it}$	Mean	$r_{it}$
Sleep	1.74	0.52	1.82	0.47	2.10	0.28
Sad or Irritable Mood	1.51	0.70	1.78	0.66	1.37	0.61
Appetite	1.04	0.28	0.99	0.35	1.41	0.33
Concentration/Decision Making	1.12	0.55	1.25	0.62	1.07	0.50
Self-View	1.05	0.67	1.27	0.65	0.99	0.52
Thoughts of Death or Suicide	0.36	0.49	0.37	0.44	0.39	0.41
General Interest	0.72	0.68	0.92	0.71	0.78	0.54
Energy Level	0.91	0.63	0.99	0.64	0.90	0.51
Restlessness/Agitation	1.00	0.45	1.14	0.48	1.19	0.53
Scale Mean	9.46		10.55		10.21	
Scale SD	5.68		5.67		5.09	
$\alpha$	0.84		0.84		0.78	

**Abbreviations:** CTT, Classical Test Theory; QIDS-A<sub>17-C</sub>, Quick Inventory of Depressive Symptomatology, Adolescent, Clinician-Report;  $r_{it}$ , Item-Total Correlation; SD, Standard Deviation; SR, Self-Report.

**Table 3** CTT Analysis of the CDRS-R

Item	(n = 103)	
	Mean	$r_{it}$
Impaired schoolwork	2.95	0.66
Difficulty having fun	2.69	0.81
Social withdrawal	2.56	0.75
Appetite disturbance	2.88	0.47
Sleep disturbance	2.20	0.64
Excessive fatigue	2.75	0.71
Physical complaints	1.95	0.50
Irritability	3.50	0.40
Excessive guilt	1.84	0.61
Low self-esteem	3.14	0.72

(Continued)

**Table 3** (Continued).

Item	(n = 103)	
	Mean	$r_{it}$
Depressed feelings	2.88	0.80
Morbid ideas	1.43	0.38
Suicidal ideas	1.72	0.58
Excessive weeping	2.21	0.68
Depressed facial affect	1.94	0.61
Listless speech	1.43	0.53
Hypoactivity	1.45	0.62
Scale Mean	39.53	
Scale SD	15.72	
$\alpha$	0.92	

**Abbreviations:** CDRS-R, Children's Depression Rating Scale-Revised; CTT, Classical Test Theory;  $r_{it}$ , Item-Total Correlation; SD, Standard Deviation.

Across QIDS-A<sub>17</sub> measures, sleep disturbance was the most commonly endorsed symptom, followed by sad or irritable mood. Appetite disturbance was also frequently endorsed on the QIDS-A<sub>17</sub>-SR. Sad/irritable mood and loss of general interest were most strongly related to overall depression across QIDS-A<sub>17</sub> measures. Symptoms least reported across QIDS-A<sub>17</sub> measures included thoughts of death or suicide and loss of general interest. Of note is that the item-total correlations of these two items were lower than the others, likely due to the lower number of participants endorsing these items.

Removal of the irritability item for comparison showed only one observable difference across the QIDS-A<sub>17</sub> measures: the sad mood item (without irritability) was endorsed slightly less frequently while relating more robustly to overall depression.

Irritability and low self-esteem were most frequently endorsed on the CDRS-R, while difficulty having fun, depressed feelings and social withdrawal were most highly related to overall depression. Morbid ideation and listless speech were endorsed least frequently. Note that symptoms with high item-total correlations (ie, strongly related to depression) on the CDRS-R had corresponding items that were highly correlated on the QIDS-A<sub>17</sub> measures.

### Correlations Among Measures

Total scores on both QIDS-A<sub>17</sub>-C measures (Adolescent and Composite) showed high correlations with the CDRS-R ( $r = 0.78$  and  $0.89$ , respectively). The QIDS-A<sub>17</sub>-SR correlated moderately with the QIDS-A<sub>17</sub>-C(Adolescent) ( $r = 0.69$ ), the QIDS-A<sub>17</sub>-C(Composite) ( $r = 0.66$ ), and the CDRS-R ( $r = 0.63$ ). The QIDS-A<sub>17</sub>-C(Composite) correlated highly with the QIDS-A<sub>17</sub>-C(Adolescent) ( $r = 0.88$ ). All correlations were significant ( $p < 0.0001$ ).

Disattenuated intercorrelations were computed by correcting the above intercorrelations for unreliability due to measurement error. This strengthened the correlation between QIDS-A<sub>17</sub>-SR and QIDS-A<sub>17</sub>-C(Adolescent) from  $0.69$  to  $0.85$ . The disattenuated correlations between the QIDS-A<sub>17</sub>-C(Composite) and two measures (QIDS-A<sub>17</sub>-C(Adolescent) and CDRS-R) exceeded  $1.0$ , reflecting essentially perfect correlations when corrected for unreliability.

### Exploratory Factor Analysis

For all four measures (the three QIDS-A<sub>17</sub> versions and the CDRS-R), the first real obtained eigenvalue far exceeded the first randomly generated eigenvalue, while the remainder of the real eigenvalues were smaller than the random

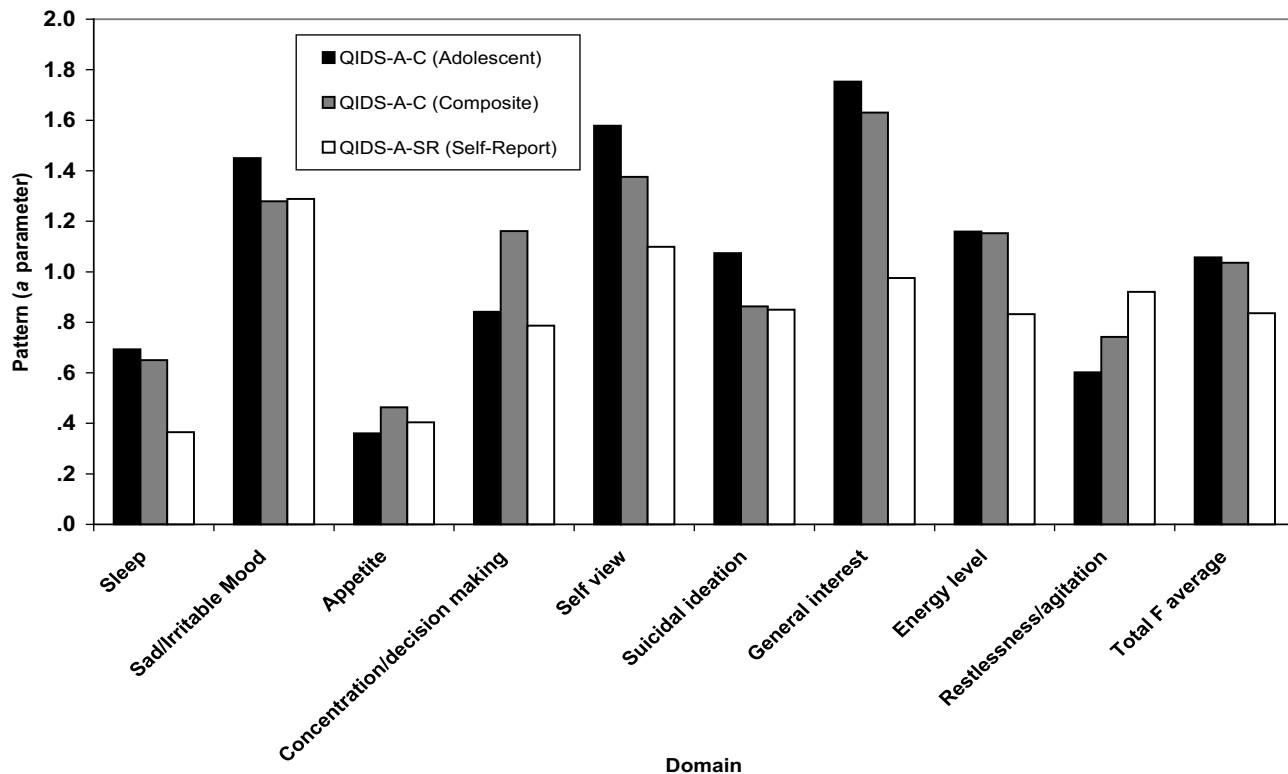
**Table 4** Dimensionality: Obtained and Simulated Eigenvalues for the QIDS-A<sub>17</sub>-C(Adolescent), QIDS-A<sub>17</sub>-C(Composite), QIDS-A<sub>17</sub>-SR, and CDRS-R

Component	Simulated Eigenvalues	Obtained Eigenvalues		
		QIDS-A <sub>17</sub> -C (Adolescent)	QIDS-A <sub>17</sub> -C (Composite)	QIDS-A <sub>17</sub> -SR
1	1.48	4.05	4.08	3.40
2	1.31	1.13	1.09	1.21
3	1.20	1.00	0.98	1.03
Component	Simulated Eigenvalues	Obtained Eigenvalues CDRS-R		
1	1.79	7.83		
2	1.62	1.38		
3	1.49	1.12		

eigenvalues. The first few obtained and simulated eigenvalues are listed in Table 4. These results indicate all three versions of the QIDS-A<sub>17</sub> as well as the CDRS-R were unidimensional for this sample.

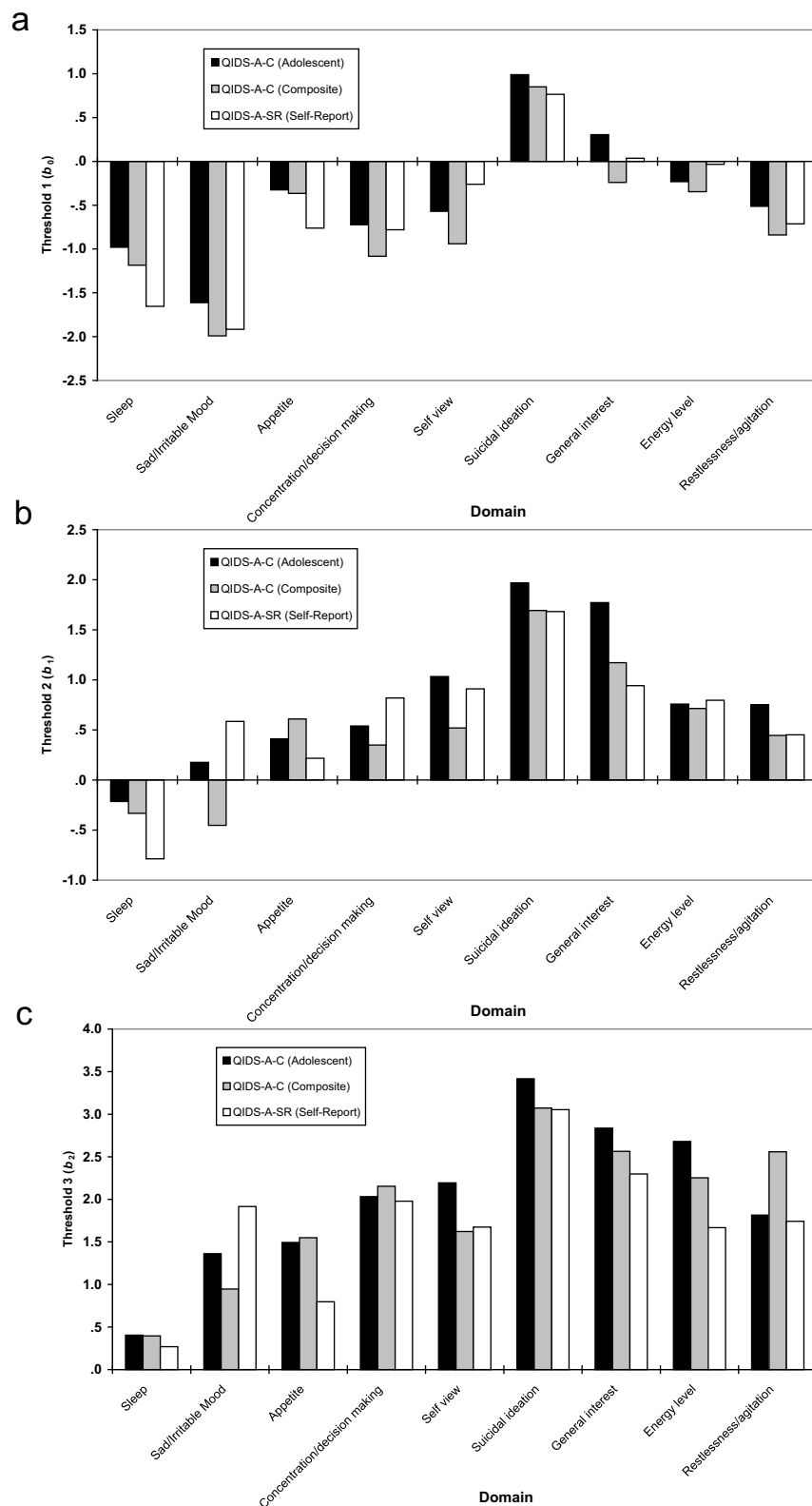
**Item Response Theory Analysis**

Figures 1 and 2a-c contain the Samejima<sup>44</sup> IRT parameter estimates for the QIDS-A<sub>17</sub>-C(Adolescent), QIDS-A<sub>17</sub>-C(Composite), and QIDS-A<sub>17</sub>-SR. For all measures, the estimates were obtained from a model in which parameters



**Figure 1** Item Response Theory Slope (a) Parameters for the QIDS-A Measures.





**Figure 2** Item Response Theory Location ( $b_0$ ,  $b_1$ ,  $b_2$ ) Parameters. (a)  $b_0$  parameters; (b)  $b_1$  parameters; (c)  $b_2$  parameters.

were allowed to vary freely. Figure 1 provides the IRT slope parameters which illustrate the pattern of influence of depression on each domain response and are similar to item-total correlations in CTT. Larger slope parameters indicate greater sensitivity to level of depression. For example, general interest, self-view, and sad/irritable mood domains were most sensitive to depression in this sample given their slope ( $a$ ) parameters are generally the largest among the domains. In contrast, appetite and sleep were the least sensitive to depression among all three QIDS-A measures, indicated by the small slope ( $a$ ) parameters.

Figure 2a–c contain the  $b_0$ ,  $b_1$ , and  $b_2$  parameter estimates for the three QIDS-A measures. These reflect the likelihood of choosing a particular response (ie, 0–3 on a QIDS-A<sub>17</sub> domain item) in each domain, regardless of how well the domain relates to depression and are similar to the item means in CTT. A higher threshold reflects a lower probability that one would choose a more pathological answer. For example, the suicidal ideation domain shows higher thresholds than the other domains, indicating the item mean for suicidal ideation is lower (ie, endorsed infrequently). Figure 2a shows lower thresholds for sad/irritable mood than for appetite, indicating people are more likely to choose a more pathological response (ie, a response of 1, 2, or 3, as opposed to 0) for sad/irritable mood than for appetite. Sleep consistently shows a higher likelihood of a pathological response (Figures 2a–c), even though it was found to be least influenced by depression (Figure 1). This pattern is similar to the CTT results, noted above, in that frequency of a domain response did not always reflect a strong relationship to depression.

Corresponding Samejima<sup>44</sup> estimates were obtained for the CDRS-R. Difficulty having fun, hypoactivity, and depressed feelings were the most discriminating items for depression, while sleep disturbance and morbid ideation were the least discriminating for the CDRS-R in this sample.

The test information function curves (Figure 3) indicate the CDRS-R was the most informative measure from about –1.5 z-score units below the mean up to 3 z-score units above the mean (ie, from mild to severe depression). The QIDS-A<sub>17</sub>-SR and QIDS-A<sub>17</sub>-C(Composite) were both slightly more informative in the range from no depression to mild

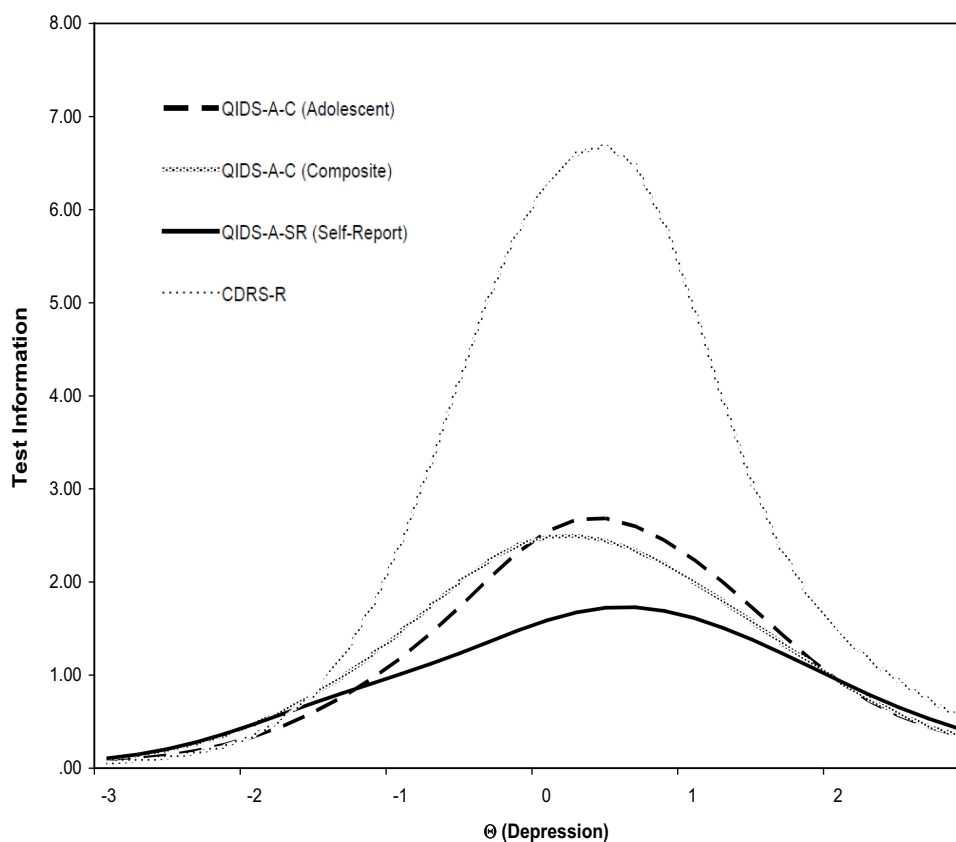


Figure 3 Test Information Function.

levels of depression (from  $\theta = -3$  to  $-1.5$ ). Among the three QIDS-A<sub>17</sub> measures, the QIDS-A<sub>17</sub>-C(Composite) performed better from  $\theta = -1.5$  to  $0$  (ie, lower to moderate levels of depression), after which the QIDS-A<sub>17</sub>-C(Adolescent) performed slightly better up to  $\theta = 1.5$ , ie, moderate depression. From this point, the QIDS-A<sub>17</sub>-SR performs just slightly better. Individually, each measure peaked at approximately  $\theta = 0.5$ , indicating all measures are more informative for moderate depression.

### Univariate Logistic Regression Analysis

The addition of each measure individually to the intercept only model to predict MDE status reduced the Likelihood Ratio chi-square by the following amounts: CDRS-R (80.59), QIDS-A<sub>17</sub>-C(Composite) (65.70), QIDS-A<sub>17</sub>-C(Adolescent) (47.20), and QIDS-A<sub>17</sub>-SR (38.85). These decreases were significant on 1 df ( $ps < 0.0001$ ). Thus, all measures were significant for detecting depression.

The regression weight, which is the log of odds ratio favoring being depressed, is the effect size estimate for univariate logistic regression. These effect sizes along with the odds ratios were examined to determine each measure's ability to differentiate between depressed versus non-depressed participants. All estimates were significant on 1 df ( $ps < 0.0001$ ) (Table 5).

### Analysis of Variance (ANOVA)

The effect size for depressed vs non-depressed participants for each measure based on a univariate ANOVA was calculated by dividing the model sum of squares by the corrected total sum of squares. The effect size values are listed in Table 5. All four measures were independently able to differentiate between the depressed and non-depressed groups at a significant level ( $ps < 0.0001$ ) (Table 5). Differences in mean scores between depressed and non-depressed participants are presented in Table 6.

**Table 5** Effect Sizes for Univariate Diagnostic Validity Analyses

Measure	Logistic Regression	Logistic Regression	ANOVA
	Log Odds Estimate	Odds Ratio	Model/Corrected Sum of Squares
QIDS-A <sub>17</sub> -C (Adolescent)	0.3258**	1.385	0.4029**
QIDS-A <sub>17</sub> -C (Composite)	0.4628**	1.588	0.5119**
QIDS-A <sub>17</sub> -SR	0.3176**	1.374	0.3351**
CDRS-R	0.2317**	1.261	0.5812**

Note: \*\*  $p < 0.0001$ .

Abbreviations: ANOVA, Analysis of Variance; CDRS-R, Children's Depression Rating Scale-Revised; QIDS-A<sub>17</sub>-C, 17-Item Quick Inventory of Depressive Symptomatology, Adolescent, Clinician-Report; SR, Self-Report.

**Table 6** Difference in Mean Scores Between Depressed and Nondepressed Participants

Measure	Depressed N = 45	Nondepressed N = 50	F-test (1, 98)	p-value
	Mean ± SD	Mean ± SD		
QIDS-A <sub>17</sub> -C(Adolescent)	13.3 ± 5.5	6.2 ± 3.3	66.13	<0.0001
QIDS-A <sub>17</sub> -C(Composite)	15.0 ± 4.5	6.8 ± 3.5	102.76	<0.0001
QIDS-A <sub>17</sub> -SR	13.4 ± 4.6	7.5 ± 3.8	49.39	<0.0001
CDRS-R	52.2 ± 13.0	20.4 ± 7.1	135.99	<0.0001

Abbreviations: CDRS-R, Children's Depression Rating Scale-Revised; SD, Standard Deviation; QIDS-A<sub>17</sub>-C, 17-Item Quick Inventory of Depressive Symptomatology, Adolescent, Clinician-Report; SR, Self-Report.

## ROC Analysis

Figure 4 displays receiver operating characteristic (ROC) curves regarding sensitivity and specificity of each scale's ability to classify depressed and non-depressed groups. Note that the CDRS-R is most sensitive to depression from a false alarm rate of 0.0 up to a false alarm rate of 0.15. At this point, the QIDS-A<sub>17</sub>-C(Composite) is most sensitive up to a false alarm rate of 0.45. The CDRS-R is once again more sensitive from this point up to a false alarm rate of 0.80, after which the curves begin to converge.

Another way to evaluate the performance of each measure is to compare the areas under each curve (c-statistic). Greater area under the curve indicates better overall performance. The areas under these curves were 0.941 (95% CI: 0.900, 0.983), 0.921 (95% CI: 0.867, 0.975), 0.854 (95% CI: 0.774, 0.935), and 0.835 (95% CI: 0.752, 0.918) for the CDRS-R, QIDS-A<sub>17</sub>-C(Composite), QIDS-A<sub>17</sub>-C(Adolescent), and QIDS-A<sub>17</sub>-SR, respectively. Thus, all measures were significantly better than chance at distinguishing MDE vs no MDE (lower limit of the CI is above the chance AUC level of 0.5). Pairwise comparison of the AUCs showed no significant differences between measures based on a Bonferroni corrected significance level of 0.0083 (all p-values above 0.0124).

Table 7 provides the thresholds, sensitivities, and specificities at four particular locations along the depression continuum for each measure. These thresholds were chosen based on scores that reflected at least 30% sensitivity (low), 50% sensitivity (medium), 70% sensitivity (high), and 90% sensitivity (very high). Such cutoff points were based on criteria used in a previous study that employed similar statistical analyses with the adult QIDS-SR<sub>16</sub>.<sup>28</sup>

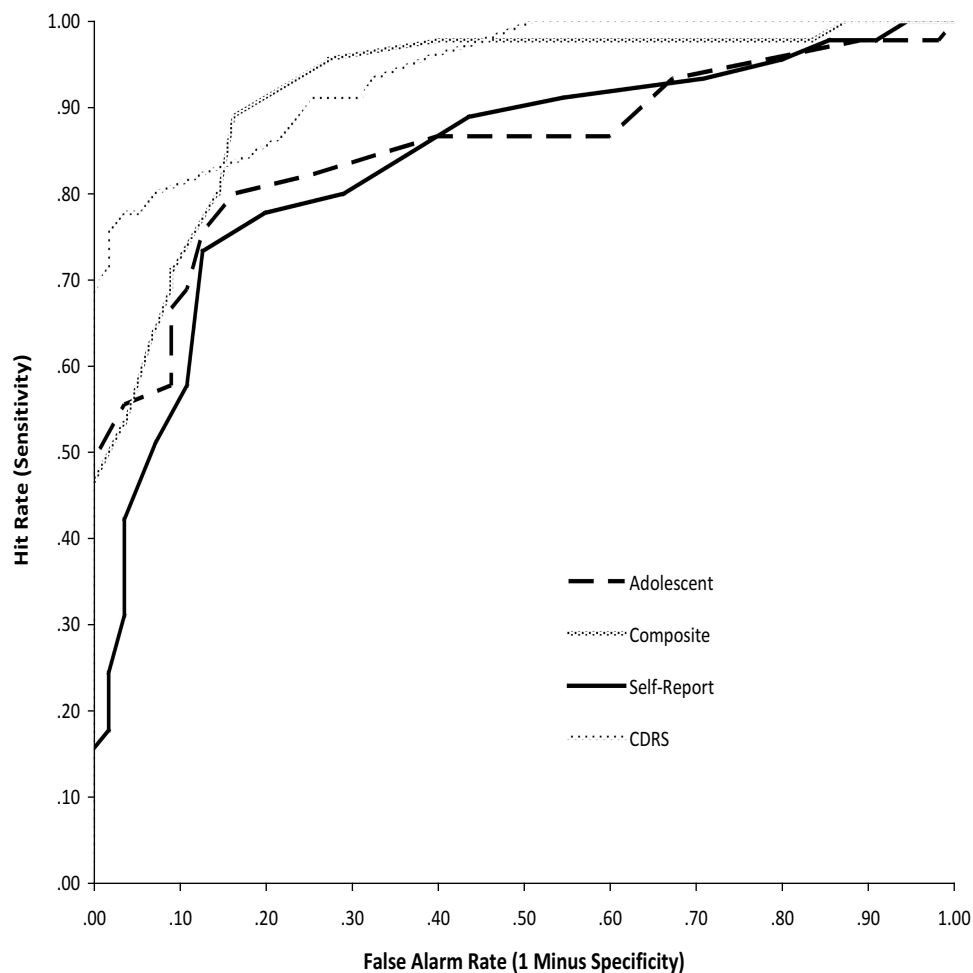


Figure 4 ROC Curves for each measure.

**Table 7** Threshold Scores, Sensitivities, and Specificities at Four Levels of Severity for the QIDS-A<sub>17</sub>-C(Adolescent), QIDS-A<sub>17</sub>-C(Composite), QIDS-A<sub>17</sub>-SR, and CDRS-R

Level	QIDS-A <sub>17</sub> -C (Adolescent)			QIDS-A <sub>17</sub> -C (Composite)			QIDS-A <sub>17</sub> -SR			CDRS-R		
	Thresh	Sens	Spec	Thresh	Sens	Spec	Thresh	Sens	Spec	Thresh	Sens	Spec
Low	5	0.38	1.00	6	0.33	1.00	6	0.31	0.96	28	0.31	1.00
Medium	8	0.56	0.96	9	0.53	0.96	8	0.51	0.93	33	0.51	1.00
High	12	0.76	0.87	12	0.71	0.91	10	0.73	0.87	40	0.71	0.98
Very High	17	0.93	0.33	15	0.96	0.73	15	0.91	0.45	49	0.91	0.75

**Abbreviations:** Thresh, threshold; Sens, sensitivity; Spec, specificity.

## Discussion

The three QIDS-A<sub>17</sub> measures (the self-report, clinician rating and the composite) and the CDRS-R (based on an integration of the parent and child clinician interviews) were of acceptable reliability and validity. Although the QIDS-A<sub>17</sub>-SR had the numerically lowest reliability, it still demonstrated strong reliability and validity and is reliable enough to be considered for use over a more time-consuming interview that combines both adolescent and parent output. The QIDS-A<sub>17</sub>-C(Adolescent), found to be slightly more reliable than the QIDS-A<sub>17</sub>-SR, could also effectively be used in place of a composite interview.

## CTT Analyses

As hypothesized, each measure demonstrated strong internal consistency in this population. The CDRS-R and both clinician versions of the QIDS-A<sub>17</sub> (Adolescent and Composite) showed the highest reliability, and coefficient alphas were comparable after accounting for the differences in test length. The QIDS-A<sub>17</sub>-SR showed slightly lower internal consistency, though well within the acceptable range.

There were similarities across the three QIDS-A<sub>17</sub> measures for tendencies to endorse symptoms, particularly sleep and symptoms that related most strongly to overall depression (sad or irritable mood and loss of general interest). The IRT analyses provided additional support regarding the reliability as well as comparability of the QIDS-A measures. Specifically, the same items shown in CTT to relate most strongly to depression (loss of general interest, sad/irritable mood) were also found in the IRT analyses to be the most discriminating for depression. The effects of removing the irritability item may indicate, in this sample, this item is also endorsed as a symptom of other disorders (eg, Bipolar Disorder, Attention-Deficit/Hyperactivity Disorder) such that it slightly decreases the loading to depression. Although irritability was the most frequently endorsed symptom on the CDRS-R, it related to overall depression only half as strongly as the depressed feelings symptom. This might further suggest, as with the QIDS-A<sub>17</sub>, irritability is an important clinical symptom of depression but not necessarily an essential factor in measuring overall depression — analogous to anxiety in depression.

Stronger correlations were found among clinician measures. This may be partly because information was gathered by a clinician and clinically-minded probes were asked to garner similarly appropriate information. Furthermore, the fact that only one clinician administered all measures to a particular participant may have increased the relationship among measures for each participant. The current study found the QIDS-A<sub>17</sub>-SR to be correlated with the QIDS-A<sub>17</sub>-C(Adolescent) and the QIDS-A<sub>17</sub>-C(Composite) (both 0.69), respectively.

## IRT Analyses

The IRT analyses provided additional support about the reliability and comparability of the QIDS-A measures. Specifically, the same items shown in CTT to relate most strongly to depression (loss of general interest, self-view, and sad/irritable mood) were also found in the IRT analyses to be the most discriminating for depression. The same was

true for items with low correlations to depression (appetite and sleep). Similarly, IRT and CTT results corresponded with items that were most frequently endorsed (ie, sleep and sad/irritable mood).

Although the three QIDS-A measures were similar in response pattern, there were some interesting differences. On the QIDS-A<sub>17</sub>-SR, sad mood was most related to depression while the clinician-rated versions found loss of general interest to correlate most with depression. This may be related to the clinician's ability to probe for pertinent information one may not think of when completing a self-report. Similarly, adolescents tended to report more pathology on self-view and general interest when interviewed by a clinician than on self-report, possibly due to querying on the clinician's part.

While the domains of the CDRS could not be directly compared to the QIDS-A measures in IRT, some similarities were found in items most influenced by depression. The depressed feelings item on the CDRS-R corresponds to the QIDS-A sad/irritable mood item, while the difficulty having fun item on the CDRS-R is somewhat similar to the loss of general interest item on the QIDS-A measures. Sleep disturbance and morbid ideation were the least discriminating on the CDRS-R. It is possible that the weak relationship between depression and morbid ideation in this sample is simply because it was not a common response, which restricts the possible threshold range necessary for accuracy in IRT.

The test information functions found all measures to be most discriminating at moderate levels of depression, with the CDRS-R the most sensitive of the measures for moderate depression. Interestingly, the QIDS-A<sub>17</sub>-SR and the QIDS-A<sub>17</sub>-C(Composite) were slightly more reliable than the other measures in detecting mild depression, at almost equal sensitivity levels. The QIDS-A<sub>17</sub>-C(Adolescent) was slightly more discriminating than the other QIDS-A<sub>17</sub> measures at a moderate level of depression.

Univariate logistic regression found all four measures significantly discriminated between depressed and non-depressed participants. The ANOVA results further substantiated all four measures significantly discriminated between depressed and non-depressed participants. In contrast to univariate logistic regression, ANOVA found the CDRS-R to be the most predictive. While this difference was minimal, it is important to note because it indicates all four measures are similar enough in predictive ability that the order varies depending on the type of analysis. As such, it is likely the order is inconsequential. Thus, each measure shows definitive discriminative validity.

ROC analysis found all measures were significantly better than change in discriminating MDE from non-MDE participants. Pairwise tests showed that no measure was significantly better than another. The CDRS-R and the QIDS-A<sub>17</sub>-C(Composite) to be the most sensitive in discriminating depression overall. This is not surprising since both measures compile data from both the adolescent and the parent, thus maximizing the information contributing to detection of depression. Although the remaining two QIDS-A measures differed slightly in sensitivity at various points, they were very similar (and robust) in overall performance, based on the area under the curve. This indicates both the QIDS-A<sub>17</sub>-C(Adolescent) and QIDS-A<sub>17</sub>-SR show a high level of accuracy in discriminating depressed patients from non-depressed patients. This supports the case that the QIDS-A<sub>17</sub>-SR is accurate enough for the purposes of correctly classifying a depressed patient. As such, the QIDS-A<sub>17</sub>-SR could be used as a simple self-report screening measure to increase detection for depression in clinic settings while minimizing time and staff burden and maximizing clinician efficiency.

The current study's diagnostic validity results regarding the QIDS-A<sub>17</sub> compare favorably to the results of a recent study that used similar analyses to compare the QIDS-SR<sub>16</sub> to two adult self-report depression rating scales in an adult population, using a structured diagnostic measure to classify depressed and non-depressed patients.<sup>50</sup> The QIDS-SR<sub>16</sub> was found to be the most valid and was recommended above the other measures for utility in private practice settings. The definitive validity of the QIDS-A<sub>17</sub> measures found in the current study supports similar utility in adolescent populations. Furthermore, the age range of the current study (ages 8 to 17) indicates the QIDS-A<sub>17</sub> can be used accurately in a wide age range, thus increasing its value. Our finding that the QIDS-A<sub>17</sub> measures showed similarly strong discriminative validity as the adult QIDS<sub>16</sub> further substantiates the similarity and utility of the QIDS measures across formats and age groups.

All four measures showed sound psychometric properties in this population. The CDRS-R and QIDS-A<sub>17</sub>-C(Composite) correlated the highest with each other and showed similar sensitivity and specificity, not surprising since both measures incorporate information from adolescent and parent. The QIDS-A<sub>17</sub>-C(Adolescent) was generally the next in overall performance, correlating highly with the QIDS-A<sub>17</sub>-C(Composite) and CDRS-R, and showing good

discrimination comparable to the QIDS-A<sub>17</sub>-C(Composite), particularly at moderate levels of depression. The strong overall performance of the QIDS-A<sub>17</sub>-C(Adolescent) and the favorable comparisons to the two composite measures (QIDS-A<sub>17</sub>-C(Composite) and CDRS-R) indicate the QIDS-A<sub>17</sub>-C(Adolescent) could sufficiently replace either the CDRS-composite or the QIDS-A composite measure.

More importantly, the QIDS-A<sub>17</sub>-SR showed satisfactory reliability, validity, and correlation with the other measures. The QIDS-A<sub>17</sub>-SR was not far from the QIDS-A<sub>17</sub>-C(Adolescent) in performance and demonstrated similar sensitivity and specificity, particularly at lower and higher levels of depression. It was not surprising the QIDS-A<sub>17</sub>-SR was slightly less reliable than the other measures since they had the advantage of clinician experience and judgment to help probe for salient information. Despite this difference, the QIDS-A<sub>17</sub>-SR consistently performed at levels indicative of a reliable and valid measure. Additionally, the QIDS-A<sub>17</sub>-SR was slightly more discriminating than the CDRS-R and the QIDS-A<sub>17</sub>-C(Adolescent) at very low levels of depression. Based on this evidence, the QIDS-A<sub>17</sub>-SR is suitable for use on its own.

It is important to consider the intended purpose of the QIDS-A<sub>17</sub> measures. The QIDS-A<sub>17</sub> is not a structured diagnostic instrument, but rather a measure of depressive symptomatology (ie, severity and frequency of symptoms) may also be used to screen for the likely presence of an MDE. As such, it is not essential QIDS-A<sub>17</sub> measures be perfectly sensitive in classifying depressed from nondepressed patients. What is important is they show good psychometric properties in measuring the symptoms of depression, which this study has demonstrated. Another implication of our study is that converting the CDRS-R to a self-report might be worth considering since the self-report of the QIDS-A was comparable to the composite clinician-rated version.

## Limitations

The sample size was relatively small. These findings should be confirmed in larger samples. Findings from this pediatric psychiatric outpatient clinic sample from a university hospital setting may not generalize to other clinical settings.

The MDE diagnosis was not obtained from a structured diagnostic interview, but rather from a physician-completed checklist of DSM-IV<sup>32</sup> symptoms required to diagnose an MDE. Furthermore, the MDE checklist and the depression rating scales were occasionally completed by the same physician, so the rater was not always blind to the diagnosis. Also, since each participant's QIDS-A<sub>17</sub>-C measures and CDRS-R were always completed by the same evaluator, the evaluator had data from both measures to guide their scores, which may have increased the correlation between these measures. In addition, we only have one time point of data and thus are not able to assess sensitivity to change.

Due to the variety of clinicians assisting on the study, the order of administration was not randomized. The order of measures could not be monitored when completed by a clinician other than the study coordinator. The youth always completed the QIDS-A<sub>17</sub>-SR first, then the clinicians interviewed the child/adolescent first to complete the QIDS-A<sub>17</sub>-C(Adolescent), followed by the parent, and then rated QIDS-A<sub>17</sub>-C(Composite) and CDRS-R.

Although every effort was made to recruit consecutive outpatients, potential participants may have been missed due to failure of the treating clinician to mention the study or the lack of response to follow-up contact when testing was scheduled at a later clinic date.

## Clinical Implications

The finding that all QIDS-A<sub>17</sub> measures demonstrated strong psychometric properties in this population indicates the QIDS-A<sub>17</sub> is an appropriate measure for depression in youth. Furthermore, results indicate that any of the QIDS-A<sub>17</sub> measures could be used in place of the CDRS-R for symptom measurement and depression screening, though the CDRS-R might be more appropriate in a precise research study as it was more sensitive to detecting slight differences in level of depression. All four measures demonstrated comparable discriminative validity, further supporting the use of the QIDS-A<sub>17</sub>. Given the QIDS-A<sub>17</sub> covers all 9 DSM-IV<sup>42</sup> symptom domains for depression, the validation of this measure meets an established need.

Although the QIDS-A<sub>17</sub>-SR was slightly less reliable than the clinician measures, it demonstrated satisfactory psychometric properties and diagnostic utility. If one balances the slight sacrifice in reliability with the need for this type of valid, time- and cost-effective tool—particularly given rising health care costs—the loss is minimal. Since it is

reliable and valid and only takes 5 to 7 minutes to complete, the QIDS-A<sub>17</sub>-SR would be particularly useful in busy clinical environments, including a pediatrician's office.

Although irritability is an important diagnostic symptom in adolescents, the inclusion of the irritability item in the QIDS-A<sub>17</sub> made little difference in the performance of the rating scales. This suggests the QIDS<sub>16</sub> versions that do not include an irritability item are not lacking in measuring depression in adolescents. It is possible the QIDS<sub>16</sub> version may be acceptable for use in adolescents, potentially eliminating the need for a separate adolescent version.

## Conclusion

The satisfactory psychometric properties of the QIDS-A<sub>17</sub>-SR or the QIDS-A<sub>17</sub>-C (Adolescent) suggest they provide more time and resource efficient, yet acceptable, options than ratings that require parental and adolescent interviews. The QIDS-A<sub>17</sub>-SR is sufficient to use in place of the clinician-rated measures where time, staffing and health care costs are issues.

## Licensing and Distribution

Licensing and distribution of the QIDS-A is managed by The University of Texas Southwestern Medical Center Office for Technology Development. At the time of publishing, the self-report and clinician-rated version of the QIDS-A are available without charge to non-commercial users. Fees may apply to commercial users, IT companies, funded academic users or healthcare organizations. Requests for information and licensing of the QIDS-A should be emailed to [TechnologyDevelopment@mednet.swmed.edu](mailto:TechnologyDevelopment@mednet.swmed.edu).

## Acknowledgments

The authors acknowledge that the primary source of this manuscript is based on Dr. Haley's unpublished dissertation.<sup>51</sup> Additionally, we are grateful to the patients, families, and clinical staff of Children's Health Outpatient Psychiatry clinic for their participation in this project.

## Disclosure

Dr. Kennard receives royalties from Guilford Press and is on the board of the Jerry M. Lewis MD Research Foundation and the George G. and Alva Hudson Smith Foundation. Dr. Kennard has research support from American Foundation for Suicide Prevention (AFSP), the National Institutes of Health, Patient-Centered Outcomes Research Institute (PCORI), and the State of Texas. Dr. Carmody is a consultant for Alkermes, Inc. Dr. Emslie is a consultant for Lundbeck and Neuronetics. Dr. Emslie has research support from American Foundation for Suicide Prevention (AFSP), Janssen Pharmaceuticals, Janssen Research & Development, the National Institutes of Health, Patient-Centered Outcomes Research Institute (PCORI), and the State of Texas. Dr. A. John Rush has received consulting fees from Compass Inc., Curbstone Consultant LLC, Emmes Corp., Evecxia Therapeutics, Inc., Holmusk Technologies, Inc., ICON, PLC, Johnson and Johnson (Janssen), Liva-Nova, MindStreet, Inc., Neurocrine Biosciences Inc., Otsuka-US; speaking fees from Liva-Nova, Johnson and Johnson (Janssen); and royalties from Wolters Kluwer Health, Guilford Press and the University of Texas Southwestern Medical Center, Dallas, TX (for the Inventory of Depressive Symptoms and its derivatives). He is also named co-inventor on two patents: US Patent No. 7,795,033: Methods to Predict the Outcome of Treatment with Antidepressant Medication, Inventors: McMahon FJ, Laje G, Manji H, Rush AJ, Paddock S, Wilson AS; and US Patent No. 7,906,283: Methods to Identify Patients at Risk of Developing Adverse Events During Treatment with Antidepressant Medication, Inventors: McMahon FJ, Laje G, Manji H, Rush AJ, Paddock S. The authors report no other conflicts of interest in this work.

## References

1. Bitsko RH, Claussen AH, Lichstein J, et al. Mental health surveillance among children - United States, 2013–2019. *MMWR Suppl.* 2022;71(2):1–42. doi:10.15585/mmwr.su7102a1
2. Costello EJ, Maughan B. Annual research review: optimal outcomes of child and adolescent mental illness. *J Child Psychol Psychiatry.* 2015;56(3):324–341. doi:10.1111/jcpp.12371



3. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. *Lancet*. 2012;379(9820):1056–1067. doi:10.1016/S0140-6736(11)60871-4
4. Thapar A, Eyre O, Patel V, Brent D. Depression in young people. *Lancet*. 2022;400(10352):617–631. doi:10.1016/S0140-6736(22)01012-1
5. Ryan ND. Treatment of depression in children and adolescents. *Lancet*. 2005;366(9489):933–940. doi:10.1016/s0140-6736(05)67321-7
6. Lebrun-Harris LA, Ghandour RM, Kogan MD, Warren MD. Five-year trends in US children's health and well-being, 2016–2020. *JAMA Pediatr*. 2022;176(7):e220056–e220056. doi:10.1001/jamapediatrics.2022.0056
7. Lu W. Adolescent Depression: national Trends, Risk Factors, and Healthcare Disparities. *Am J Health Behav*. 2019;43(1):181–194. doi:10.5993/ajhb.43.1.15
8. Merikangas KR, He JP, Burstein M, et al. Service utilization for lifetime mental disorders in US adolescents: results of the National Comorbidity Survey-Adolescent Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry*. 2011;50(1):32–45. doi:10.1016/j.jaac.2010.10.006
9. Brent D, Emslie G, Clarke G, et al. Switching to another SSRI or to venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression: the TORDIA randomized controlled trial. *JAMA*. 2008;299(8):901–913. doi:10.1001/jama.299.8.901
10. Dwyer JB, Stringaris A, Brent DA, Bloch MH. Annual research review: defining and treating pediatric treatment-resistant depression. *J Child Psychol Psychiatry*. 2020;61(3):312–332. doi:10.1111/jcpp.13202
11. Abright AR, Grudnikoff E. Measurement-based care in the treatment of adolescent depression. *Child Adolesc Psychiatr Clin N Am*. 2020;29(4):631–643. doi:10.1016/j.chc.2020.06.003
12. Liu FF, Adrian MC. Is treatment working? Detecting real change in the treatment of child and adolescent depression. *J Am Acad Child Adolesc Psychiatry*. 2019;58(12):1157–1164. doi:10.1016/j.jaac.2019.02.011
13. Jeffrey J, Klomhaus A, Enenbach M, Lester P, Krishna R. Self-report rating scales to guide measurement-based care in child and adolescent psychiatry. *Child Adolesc Psychiatr Clin N Am*. 2020;29(4):601–629. doi:10.1016/j.chc.2020.06.002
14. Fortney JC, Unützer J, Wrenn G, et al. A tipping point for measurement-based care. *Psychiatr Serv*. 2017;68(2):179–188. doi:10.1176/appi.ps.201500439
15. Bickman L, Kelley SD, Breda C, de Andrade AR, Riemer M. Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. *Psychiatr Serv*. 2011;62(12):1423–1429. doi:10.1176/appi.ps.002052011
16. Poznanski EO, Cook SC, Carroll BJ. A depression rating scale for children. *Pediatrics*. 1979;64(4):442–450. doi:10.1542/peds.64.4.442
17. Poznanski EO, Mokros H. Children's depression rating scale-revised. *Psychopharmacol Bull*. 1985;21(4):979–989.
18. Poznanski EO, Mokros HB. *Children's Depression Rating Scale, Revised (CDRS-R) Manual*. Western Psychological Services; 1996.
19. Cheung AH, Emslie GJ, Mayes TL. Review of the efficacy and safety of antidepressants in youth depression. *J Child Psychol Psychiatry*. 2005;46(7):735–754. doi:10.1111/j.1469-7610.2005.01467.x
20. Brooks SJ, Kutcher S. Diagnosis and measurement of adolescent depression: a review of commonly utilized instruments. *J Child Adolesc Psychopharmacol*. 2001;11(4):341–376. doi:10.1089/104454601317261546
21. Myers K, Winters NC. Ten-year review of rating scales. II: scales for internalizing disorders. *J Am Acad Child Adolesc Psychiatry*. 2002;41(6):634–659. doi:10.1097/00004583-200206000-00004
22. Treatment for Adolescents with Depression Study Team. Treatment for Adolescents With Depression Study (TADS): rationale, design, and methods. *J Am Acad Child Adolesc Psychiatry*. 2003;42(5):531–542. doi:10.1097/01.Chi.0000046839.90931.0d
23. Baumgartner N, Häberling I, Emery S, et al. When parents and children disagree: informant discrepancies in reports of depressive symptoms in clinical interviews. *J Affect Disord*. 2020;272:223–230. doi:10.1016/j.jad.2020.04.008
24. Rush AJ, Kraemer HC, Sackeim HA, et al. Report by the ACNP Task Force on response and remission in major depressive disorder. *Neuropsychopharmacology*. 2006;31(9):1841–1853. doi:10.1038/sj.npp.1301131
25. Mintz J, Mintz LI, Arruda MJ, Hwang SS. Treatments of depression and the functional capacity to work. *Arch Gen Psychiatry*. 1992;49(10):761–768. doi:10.1001/archpsyc.1992.01820100005001
26. Herjanic B, Reich W. Development of a structured psychiatric interview for children: agreement between child and parent on individual symptoms. *J Abnorm Child Psychol*. 1982;10(3):307–324. doi:10.1007/BF00912324
27. Wu P, Hoven CW, Bird HR, et al. Depressive and disruptive disorders and mental health service utilization in children and adolescents. *J Am Acad Child Adolesc Psychiatry*. 1999;38(9):1081–90; discussion 1090–2. doi:10.1097/00004583-199909000-00010
28. Bernstein IH, Rush AJ, Carmody TJ, Woo A, Trivedi MH. Clinical vs. self-report versions of the quick inventory of depressive symptomatology in a public sector sample. *J Psychiatr Res*. 2007;41(3–4):239–246. doi:10.1016/j.jpsychires.2006.04.001
29. 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(6):493–501. doi:10.1016/j.biopsych.2005.08.022
30. Rush AJ, Trivedi MH, Carmody TJ, et al. Self-reported depressive symptom measures: sensitivity to detecting change in a randomized, controlled trial of chronically depressed, nonpsychotic outpatients. *Neuropsychopharmacology*. 2005;30(2):405–416. doi:10.1038/sj.npp.1300614
31. 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(5):573–583. doi:10.1016/s0006-3223(02)01866-8
32. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th, text rev ed. American Psychiatric Association; 2013.
33. Bernstein IH, Rush AJ, Trivedi MH, et al. Psychometric properties of the Quick Inventory of Depressive Symptomatology in adolescents. *Int J Methods Psychiatr Res*. 2010;19(4):185–194. doi:10.1002/mpr.321
34. Zhang WY, Zhao YJ, Zhang Y, et al. Psychometric properties of the Quick Inventory of Depressive Symptomatology-Self-Report (QIDS-SR) in depressed adolescents. *Front Psychiatry*. 2020;11:598609. doi:10.3389/fpsy.2020.598609
35. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychol Bull*. 1987;101(2):213–232. doi:10.1037/0033-2909.101.2.213
36. Kovacs M. *Children's Depression Inventory Manual*. Multi-Health Systems; 1992.
37. Mayes TL, Bernstein IH, Haley CL, Kennard BD, Emslie GJ. Psychometric properties of the children's depression rating scale-revised in adolescents. *J Child Adolesc Psychopharmacol*. 2010;20(6):513–516. doi:10.1089/cap.2010.0063

38. Johnson JG, Harris ES, Spitzer RL, Williams JB. The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *J Adolesc Health*. 2002;30(3):196–204. doi:10.1016/s1054-139x(01)00333-0
39. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961;4:561–571. doi:10.1001/archpsyc.1961.01710120031004
40. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56–62. doi:10.1136/jnnp.23.1.56
41. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606–613. doi:10.1046/j.1525-1497.2001.016009606.x
42. Bell CC. DSM-IV: diagnostic and statistical manual of mental disorders. *JAMA*. 1994;272(10):828–829. doi:10.1001/jama.1994.03520100096046
43. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297–334. doi:10.1007/BF02310555
44. Samejima F. Graded response model. In: *Handbook of Modern Item Response Theory*. Vol. 1. Springer-Verlag; 1997:8.
45. Humphreys LG, Montanelli RG. An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behav Res*. 1975;10(2):193–205. doi:10.1207/s15327906mbr1002\_5
46. Portney LG, Watkins MP. Statistical measures of validity. In: *Foundations of Clinical Research: Applications to Practice*. 3 ed. McGraw-Hill Education; 2017.
47. Kraemer HC. *Evaluating Medical Tests: Objective and Quantitative Guidelines*. Vol. 26. Newbury Park, CA: Sage publications; 1992.
48. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845. doi:10.2307/2531595
49. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. McGraw-Hill; 1994.
50. Bernstein IH, Wendt B, Nasr SJ, Rush AJ. Screening for major depression in private practice. *J Psychiatr Pract*. 2009;15(2):87–94. doi:10.1097/01.pra.0000348361.03925.b3
51. Haley CL. *Improving Depressive Symptom Measurement in Adolescents: A Psychometric Evaluation of the Quick Inventory of Depressive Symptomatology, Adolescent Version (QIDS-A<sub>17</sub>)* [Doctoral dissertation]. Dallas, TX: University of Texas Southwestern Medical Center; 2009.

## Neuropsychiatric Disease and Treatment

Dovepress

### Publish your work in this journal

Neuropsychiatric Disease and Treatment is an international, peer-reviewed journal of clinical therapeutics and pharmacology focusing on concise rapid reporting of clinical or pre-clinical studies on a range of neuropsychiatric and neurological disorders. This journal is indexed on PubMed Central, the 'PsycINFO' database and CAS, and is the official journal of The International Neuropsychiatric Association (INA). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/neuropsychiatric-disease-and-treatment-journal>