A Comparison of Depressive Symptom Self-Reported Measures in the Texas Youth Depression and Suicide Research Network (TX-YDSRN)

Karabi Nandy, PhD; A. John Rush, MD; Thomas Carmody, PhD; Beth D. Kennard, PsyD; Graham J. Emslie, MD; Holli Slater, PhD; Taryn L. Mayes, MS; Melissa DeFilippis, MD; Cynthia Garza, PsyD; Eric A. Storch, PhD; Sarah M. Wakefield, MD; and Madhukar H. Trivedi, MD

Abstract

Objective: To evaluate psychometrically and provide crosswalks between 3 self-report measures of depressive symptomatology in youth in psychiatric care settings. Ratings included the Patient Health Questionnaire for Adolescents (PHQ-A), a widely used 9-item self-report; the 16-item Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}); and the 5-item Very Quick Inventory of Depressive Symptomatology–Self-Report (VQIDS-SR_5), a recent effort to create a bridge from the QIDS-SR_{16} to clinical practice.

Methods: Data from the Texas Youth Depression and Suicide Research Network Registry (August 26, 2020–May 11, 2022) were included in this work. At first visit, 795 depressed or suicidal adolescent (12–20 years of age) psychiatric outpatients completed the PHQ-A, QIDS-SR_{16}, and VQIDS-SR_5. Classical test theory and item-response theory (IRT) analyses were conducted. Crosswalks among total scales were created. Sensitivity to change over 1-month follow-up was assessed for all 3 scales (n = 682).

Results: Cronbach alphas were 0.86 (PHQ-A), 0.80 (QIDS-SR_{16}), and 0.76 (VQIDS-SR_5). Item total correlations were 0.49–0.72, 0.29–0.64, and 0.43–0.61, respectively. All 3 scales were unidimensional and sensitive to change over a 1-month period. IRT analyses revealed satisfactory item performance. Modest but significant associations were found between baseline to 1-month changes in PHQ-A and VQIDS-SR_5 total scores ($r = 0.50$, $P < .0001$) and between PHQ-A and QIDS-SR_{16} total scores ($r = 0.56$; $P < .0001$). Categorical thresholds of severity (ie, mild, moderate, severe, and very severe) were comparable between PHQ-A and QIDS-SR_{16}.

Conclusions: The PHQ-A, QIDS-SR_{16}, and VQIDS-SR_5 are unidimensional, psychometrically acceptable self-reports of depressive prevalence or severity in adolescents and young adults in this sample. Total scale scores on any measure can be converted reliably to those on any other.

J Clin Psychiatry 2024;85(1):23m14861

Author affiliations are listed at the end of this article.

Major depressive disorder (MDD) is a heterogeneous, prevalent syndrome that affects persons of all ages, including 3% of 3-to 17-year-olds.[1] In youth, MDD is often associated with substantial functional impairment and mortality.[2–5] Antidepressant medications and evidence-based psychotherapies have demonstrated efficacy in clinical research trials.[6] Suboptimal outcomes in practice, however, can result from underrecognition and undertreatment.[7–9]

Depressive symptom rating scales such as the Children's Depression Rating Scale (CDRS)[10] or the Patient Health Questionnaire (PHQ),[11] when used as screening instruments, can address underrecognition. Undertreatment has been addressed in research studies with frequent visits and diligent dose escalation informed by the regular clinical ratings of depressive symptoms with global or itemized depressive symptom measures such as the CDRS, Montgomery-Åsberg Depression Rating Scale,[12] or Hamilton Depression Rating Scale.[13] In practice, undertreatment has been addressed with less time-consuming self-reported depression ratings (Patient Reported Outcomes [PROs])[14] to inform clinical decision-making or to implement measurement-based care (MBC).[15–21] Recent practice guidelines suggest that MBC is underutilized in practice, despite evidence for its effectiveness.[6]

Scan Now See supplementary material for this article at Psychiatrist.com

Editor's Note
We encourage authors to submit papers for consideration as a part of our Focus on Childhood and Adolescent Mental Health section. Please contact Karen D. Wagner, MD, PhD, at kwagner r@psychiatrist.com.

Clinical Points

- Three patient-reported outcomes—Patient Health Questionnaire for Adolescents (PHQ-A), Quick Inventory of Depressive Symptomatology–Self-Report ($QIDS\text{-}SR_{16}$), and Very Quick Inventory of Depressive Symptomatology–Self-Report ($VQIDS\text{-}SR_5$)—can be used to assess depressive symptom severity and change over time in 8- to 20-year-olds.
- PHQ-A and $QIDS\text{-}SR_{16}$ total scores are approximately the same.
- Severity thresholds on PHQ-A are approximately matched to $VQIDS\text{-}SR_5$ total scores of 0–1 (none), 2–4 (mild), 5–7 (moderate), 8–10 (severe), and 11+ (very severe).

Practitioners rely on PROs to assess outcomes or implement MBC, while researchers often report PROs as secondary outcomes, relying on clinician-rated scales for primary outcomes. Research findings may be more readily understood and applied in practice sooner if they could be converted into metrics that clinicians use.[22] The linkage of different instruments that measure the same construct to a common metric by cocalibrating item parameters is known as a crosswalk. Indeed, some studies with adult depressed patients have provided crosswalks between total scores on clinical and self-reported ratings.[23] Crosswalks between total scores on various PROs should help clinicians, services researchers, and care system managers compile screening and outcome data collected with various PROs from different care systems. Crosswalks between standard and even briefer PROs are also being developed to facilitate their use in smartphones to more frequently gather clinical information. Frequent measurements may be necessary to assess outcomes of difficult-to-treat depression,[24,25] in which symptoms are expected to wax and wane over time. For example, the brief 5-item Very Quick Inventory of Depressive Symptomatology-Self-Report ($VQIDS\text{-}SR_5$)[26,27] was developed to provide a convenient clinical tool that could be crosswalked back to the 16-item self-reported Quick Inventory of Depressive Symptomatology ($QIDS\text{-}SR_{16}$),[25] which itself is a subset of the Inventory of Depressive Symptomatology[15] (Supplementary Table 1).

To address some of the knowledge gaps among several PROs, this report psychometrically compared and developed crosswalks among 3 patient-reported assessments measuring depressive symptoms in a sample of adolescents and young adults from the Texas Youth Depression and Suicide Research Network (TX-YDSRN).[28] Specifically, it evaluates and compares the PHQ-A[29] (the adolescent version of the 9-item PHQ), $QIDS\text{-}SR_{16}$,[18,19] and $VQIDS\text{-}SR_5$.[25,26] We were interested in whether the brief $VQIDS\text{-}SR_5$ can provide a reasonable bridge between the longer $QIDS\text{-}SR_{16}$, which is more commonly used in research, and PHQ-A, a standard in clinical practice.

Specifically, the following research questions were addressed:

1. Do these 3 scales have comparable measures of internal consistency?
2. Are these scales unidimensional, measuring the same trait?
3. How do individual items perform in relation to the overall trait in each scale?
4. Can the scale total scores be linked, so that the total score on each scale can be converted to those of the others?
5. How do the 3 scales compare to each other in detecting change between baseline and 1-month follow-up assessments?

METHODS

Study Design and Participants

Launched in August 2020, the TX-YDSRN consists of 12 academic medical institutions in Texas that contribute to a registry of children and adolescents (ages 8 to 20 years) with a positive screening for depression or suicidal ideation or behavior or who are actively receiving treatment for depression at participating clinics. Ineligibility criteria included having active psychotic symptoms or acute medical or psychological condition(s) that would make participation difficult or unsafe. We did not assess personality disorder or autism spectrum disorder diagnoses as part of this study. We also did not test for IQ, although if there were concerns about the cognitive ability of a participant to complete the forms/measures, they were excluded. Details about the TX-YDSRN and further characterization of the sample can be found elsewhere.[28] This report focuses exclusively on depressed adolescents, ages 12–20 (n = 795), who were enrolled and completed their baseline visit between August 26, 2020, and May 11, 2022.

Measures

Demographic characteristics (ie, age, race, Hispanic ethnicity, and sex assigned at birth) were based on self-report.

Patient Health Questionnaire for Adolescents (PHQ-A). The PHQ-A is a 9-item self-report questionnaire[29] that measures the past-2-weeks prevalence of each of the 9 criterion symptoms that define a major depressive episode according to *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (*DSM-5*),[30] with rephrasing of items from the adult PHQ[11] to better suit adolescents. This includes adding irritation to the melancholy mood item 1 and adding schooling to the focus item 7.[29] Each item is rated 0–3: 0 ("Not at all"), 1 ("Several days"), 2 ("More than half the days"), and 3 ("Nearly every day"), with total score ranging from 0 to 27 and higher scores indicating more severe depressive symptoms (Supplementary Table 1).

Table 1.

Sample Characteristics at Baseline (n = 795)

Demographic characteristics	Value
Sex (assigned at birth), n (%)	
Male	200 (25.16)
Female	595 (74.84)
Race, n (%)	
White	524 (65.91)
African-American	85 (10.69)
Others	86 (10.82)
More than 1 race	84 (10.57)
Unknown	16 (2.01)
Hispanic, n (%)	
Yes	427 (54.67)
No	354 (46.33)
Age, n (%)	
12–14 y	236 (29.69)
15–17 y	351 (44.15)
18–20 y	208 (26.16)
Age, mean (SD), y	15.90 (2.28)
Mental health characteristics, mean (SD)	
PHQ-A (range: 0–27)	12.92 (6.32)
QIDS-SR_{16} (range: 0–27)	13.29 (5.30)
VQIDS-SR_5 (range: 0–15)	6.50 (3.46)
Depressive symptoms, n (%)[a]	
Minimal	83 (10.44)
Mild	167 (21.01)
Moderate	221 (27.80)
Severe	192 (24.15)
Very severe	132 (16.60)

[a]Categories based on PHQ-A: 0–4, 5–9, 10–14, 15–19, and 20–27 are considered as minimal, mild, moderate, severe, and very severe symptom severity, respectively (see Johnson et al[29]).

Abbreviations: PHQ-A = Patient Health Questionnaire for Adolescents, QIDS-SR_{16} = 16-item Quick Inventory of Depressive Symptomatology–Self-Report, VQIDS-SR_5 = 5-item Very Quick Inventory of Depressive Symptomatology–Self-Report.

Table 2.

Item-Total Correlations for PHQ-A, QIDS-SR_{16}, and VQIDS-SR_5 at Baseline (n = 795)

	Corrected item-total correlation
PHQ-A	
Item 1: Little interest or loss of pleasure	0.72
Item 2: Feeling down, depressed, or hopeless	0.67
Item 3: Trouble falling or staying asleep, or sleeping too much	0.54
Item 4: Feeling tired or having little energy	0.53
Item 5: Poor appetite or overeating	0.63
Item 6: Feeling bad about yourself—or that you are a failure	0.63
Item 7: Trouble concentrating	0.57
Item 8: Moving or speaking so slowly that other people could notice	0.49
Item 9: Thoughts that you would be better off dead or of hurting yourself	0.56
QIDS-SR_{16}	
Item 1: Involvement (general interest)	0.58
Item 2: Mood (sad)	0.60
Item 3: Sleep	0.29
Item 4: Energy/fatigability	0.64
Item 5: Appetite/weight	0.37
Item 6: Outlook (self)	0.49
Item 7: Concentration/decision making	0.54
Item 8: Psychomotor	0.46
Item 9: Suicidal ideation (thoughts of death or suicide)	0.51
VQIDS-SR_5	
Item 1: Involvement (general interest)	0.55
Item 2: Mood (sad)	0.58
Item 3: Energy/fatigability	0.61
Item 4: Outlook (self)	0.48
Item 5: Psychomotor	0.43

Abbreviations: PHQ-A = Patient Health Questionnaire for Adolescents, QIDS-SR_{16} = 16-item Quick Inventory of Depressive Symptomatology–Self-Report, VQIDS-SR_5 = 5-item Very Quick Inventory of Depressive Symptomatology–Self-Report.

16-item Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR_{16}). The 16-item QIDS-SR_{16}[18,19] was derived from the 30-item IDS-SR,[15,16] which was developed as a measure of depressive symptom severity over the prior 7 days. The 16 items on the QIDS-SR_{16} were selected from the 30-item Inventory of Depressive Symptomatology Self-Report (IDS-SR) items to assess 9 *DSM-IV* and *DSM-5* criterion symptom domains, which are identical to the 9 domains assessed by PHQ-9 (ie, sad mood; concentration; self-criticism; suicidal ideation; interest; energy/fatigue; sleep disturbance [initial, middle, and late insomnia or hypersomnia]; decrease/increase in appetite/weight; and psychomotor agitation/retardation). The total score on the QIDS-SR_{16} ranges from 0 to 27 with higher scores indicating more severe depressive symptoms. The QIDS-SR_{16} is sensitive to change.[18,19,23]

5-item Very Quick Inventory of Depressive Symptomatology–Self-Report (VQIDS-SR_5). VQIDS-SR_5 was created out of the QIDS-SR_{16} as a short, 5-item measure of the core symptoms of depression[25,26] that best match 5 of the 6 items on the 6-item Hamilton rating scale,[31–33] except for anxiety (Supplementary Table 1). The range is from 0 to 15, with higher scores indicating greater depressive symptom severity. It is sensitive to change in adults,[26] and crosswalks in adults have been reported.[25]

Mini-International Neuropsychiatric Interview (MINI)-Kid. The MINI is a structured psychiatric interview that assesses the 30 most common disorders in pediatric mental health.[34] In this study, we focused on the "primary diagnosis" using the MINI item, "Which problem troubles you the most or dominates the others or came first in the natural history?" which lists 37 different primary diagnoses. Note that for this question, youth and guardian (if youth was under 18 years old) were interviewed, and the clinician made the assessment of the primary diagnosis. From the responses to this item, we took the following as indicators of depression: major depressive disorder (past 2 weeks/past/recurrent),

Table 3.

IRT Parameter Estimates for PHQ-A, QIDS-SR$_{16}$, and VQIDS-SR$_5$ (n = 795)

IRT parameters	a (Slope)	b_0 (Threshold)[a]	b_1 (Threshold)[a]	b_2 (Threshold)[a]
PHQ-A				
Item 1: Little interest or loss of pleasure	3.08	−1.28	0.17	0.96
Item 2: Feeling down, depressed, or hopeless	2.34	−1.22	0.22	1.12
Item 3: Trouble falling or staying asleep, or sleeping too much	1.35	−2.03	−0.48	0.56
Item 4: Feeling tired or having little energy	1.34	−1.13	0.17	1.09
Item 5: Poor appetite or overeating	1.96	−1.62	−0.37	0.62
Item 6: Feeling bad about yourself—or that you are a failure	2.11	−1.04	0.18	0.99
Item 7: Trouble concentrating	1.42	−1.45	−0.26	0.78
Item 8: Moving or speaking so slowly that other people could notice	1.25	−0.30	0.92	1.90
Item 9: Thoughts that you would be better off dead or of hurting yourself	2.03	0.21	1.24	1.88
QIDS-SR$_{16}$				
Item 1: Involvement (general interest)	1.86	−0.76	0.57	1.77
Item 2: Mood (sad)	1.90	−1.63	−0.08	1.16
Item 3: Sleep	0.69	−5.85	−3.35	−0.27
Item 4: Energy/fatigability	2.14	−1.17	0.10	1.32
Item 5: Appetite/weight	0.85	−1.52	0.53	1.95
Item 6: Outlook (self)	1.35	−1.25	−0.06	0.90
Item 7: Concentration/decision making	1.53	−1.39	−0.06	1.87
Item 8: Psychomotor	1.24	−1.85	0.42	1.59
Item 9: Suicidal ideation (thoughts of death or suicide)	1.51	−0.18	1.24	2.40
VQIDS-SR$_5$				
Item 1: Involvement (general interest)	1.87	−0.76	0.56	1.76
Item 2: Mood (sad)	1.81	−1.67	−0.08	1.19
Item 3: Energy/fatigability	2.29	−1.15	0.10	1.29
Item 4: Outlook (self)	1.30	−1.29	−0.07	0.92
Item 5: Psychomotor	1.40	0.22	1.40	2.70

[a]b_0 represents the level of depression at which a subject would be equally likely to endorse a rating of 0 versus 1, 2, or 3; b_1 represents equal likelihood of 0 or 1 versus 2 or 3; b_2 represents equal likelihood of 0, 1, or 2 versus 3.
Abbreviations: IRT = item-response theory, PHQ-A = Patient Health Questionnaire for Adolescents, QIDS-SR$_{16}$ = 16-item Quick Inventory of Depressive Symptomatology–Self-Report, VQIDS-SR$_5$ = 5-item Very Quick Inventory of Depressive Symptomatology–Self-Report.

persistent depressive disorder (current), bipolar I disorder (current and/or past), bipolar II disorder (current and/or past), other specified bipolar and related disorders (current and/or past), bipolar I disorder with psychotic features (current and/or past), and major depressive disorder with psychotic features (current and/or past).

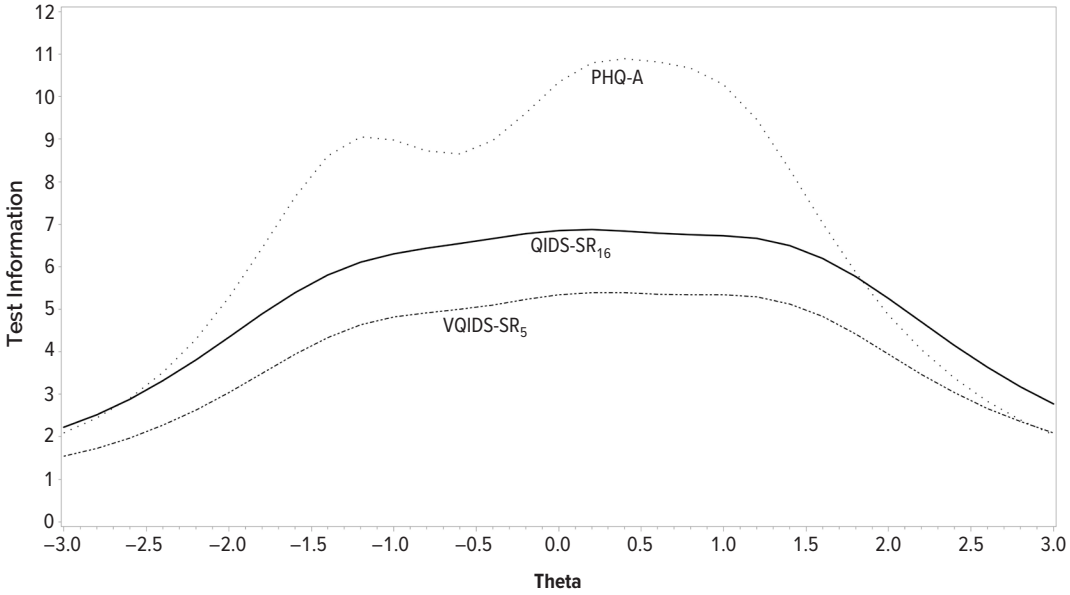
Statistical Data Analyses

Classical test theory was used to assess the internal consistency of each of the 3 scales. Corrected item-total correlations for each item within a scale were calculated. Values > 0.3 indicate that an item discriminates well.[35] Next, parallel analysis[36] assessed the unidimensionality of each scale. If established, unidimensionality indicates that all items of a scale adhere to a single latent trait and that a "total score" based on summing scale items is a valid representation of that latent trait.

Next, to assess how individual items within each scale perform in relation to the scale's overall trait, a graded response model,[37] which is appropriate for scales with multicategory items such as the QIDS-SR$_{16}$, VQIDS-SR$_5$, and PHQ-A, was implemented using MULTILOG.[38] In the Samejima model, the "a" parameter represents the ability of the item to distinguish between levels of depression and is closely related to the correlation between the item and total score. The "b" parameters represent locations on the item-response theory (IRT) depression severity scale (theta). Theta is an IRT-based measure of depression severity in standard deviation units where 0 represents average depression severity. Each QIDS-SR$_{16}$, VQIDS-SR$_5$, and PHQ-A item has 4 levels which result in 3 location parameters (b_0, b_1, and b_2). The parameter b_0 represents the level of depression at which a subject would be equally likely to endorse 0 versus 1, 2, or 3, while b_1 represents equal likelihood of 0 or 1 versus 2 or 3 and b_2 represents equal likelihood of 0, 1, or 2 versus 3. Higher values of "b" parameters indicate that higher levels of depression are needed for a subject to endorse a given level of that item and imply that the item is less frequently endorsed.

The next goal was to determine whether the total score on each scale could be linked and thus converted to that of the others. Conversion tables were constructed that equate

Figure 1.
Test Information Functions for the PHQ-A, QIDS-SR$_{16}$, and VQIDS-SR$_5$



Abbreviations: PHQ-A = Patient Health Questionnaire for Adolescents, QIDS-SR$_{16}$ = 16-item Quick Inventory of Depressive Symptomatology–Self-Report, VQIDS-SR$_5$ = 5-item Very Quick Inventory of Depressive Symptomatology–Self-Report.

total scores for each pair of scales with equivalent levels of depression severity by computing the IRT score (theta) for each total score of each scale, using the method of Orlando et al[39] and equating total scores with the closest IRT scores. The graded IRT model was also used to compute the test information function (TIF) for each scale.[40] The "test information" provided by a scale is the inverse of the standard error of the total score of the scale. A total score that provides a more precise estimate of symptom severity contains more "information" than a less precise estimate.

Finally, we assessed whether the 3 scales were sensitive to change by (1) calculating percent change between baseline and 1-month follow-up for each total score; (2) calculating the correlation coefficients of these quantities between pairs of scales; and (3) calculating effect sizes for the total score and items in each scale.

RESULTS

The sample consisted of all youth (ages 12–20 years) who completed the PHQ-A and IDS-SR during their baseline visit and had depression as their primary diagnosis based on the MINI (n = 795). Baseline data were used to answer research questions 1–4. For research question 5, 1-month post-baseline visit data were used in addition to the baseline data. Out of the n = 795 eligible youth at baseline, the majority (n = 682) had IDS-SR and PHQ-A data at 1-month follow-up visit. The 113 youth without follow-up data did not differ from those who did by demographic variables (ie, sex, age, race, and ethnicity).

Sample Characteristics

The majority of the analytic sample was White (n = 524, 65.9%), female (n = 595, 74.8%), and Hispanic (n = 427, 54.7%) (Table 1). Distributions of QIDS-SR$_{16}$ and PHQ-A total scores by domain at baseline are in Supplementary Figure 1.

Classical Test Theory Findings

Standardized Cronbach alphas were 0.86 (PHQ-A), 0.80 (QIDS-SR$_{16}$), and 0.76 (VQIDS-SR$_5$). Additionally, the corrected item-total correlation for each scale at baseline varied between 0.49–0.72 (PHQ-A), 0.29–0.64 (QIDS-SR$_{16}$), and 0.43–0.61 (VQIDS-SR$_5$) (Table 2).

Dimensionality

For PHQ-A, the first eigenvalue from the sample data (4.35) was larger than the first eigenvalue of the simulated data (1.16), and the second eigenvalue from the sample data was lower than the second eigenvalue of the simulated data (0.996 vs 1.10), demonstrating its unidimensionality. Similar findings were observed for QIDS-SR$_{16}$ (sample vs simulated eigenvalues: 3.69 vs 1.16; 0.98 vs 1.10) and VQIDS-SR$_5$ (2.59 vs 1.10; 0.81 vs 1.04).

Item-Response Theory Findings

Table 3 presents the IRT item parameters of the 3 scales. For the QIDS-SR$_{16}$ and VQIDS-SR$_5$, the item related to energy/fatigability was most sensitive in distinguishing levels of depression, while "Little interest or loss of pleasure" was most sensitive for the PHQ-A. For QIDS-SR$_{16}$ and PHQ-A, the suicide ideation item required the

Table 4.

Conversion of VQIDS-SR$_5$ to QIDS-SR$_{16}$ and VQIDS-SR$_5$ to PHQ-A

Conversion of VQIDS-SR$_5$ to QIDS-SR$_{16}$		Conversion of VQIDS-SR$_5$ to PHQ-A		Conversion between QIDS-SR$_{16}$ and PHQ-A	
VQIDS-SR$_5$	QIDS-SR$_{16}$	VQIDS-SR$_5$	PHQ-A	QIDS-SR$_{16}$	PHQ-A
0	0–3	0	0–1	0–1	0
1	4–5	1	2–3	2–3	1
2	6–7	2	4–5	4	2
3	8	3	6–7	5	3
4	9–10	4	8–9	6	4–5
5	11	5	10–11	7	6
6	12–13	6	12–13	8	7
7	14	7	14	9	8
8	15–16	8	15–16	10	9
9	17	9	17–18	11	10
10	18–19	10	19	12	11
11	20	11	20–21	13	12–13
12	21–22	12	22–23	14	14
13	23	13	24–25	15	15
14	24–25	14	26	16	16
				17	17
				18	18–19
				19	20
				20	21
				21	22
				22	23
				23	24
				24	25
				25	26
				26–27	27

Abbreviations: PHQ-A = Patient Health Questionnaire for Adolescents, QIDS-SR$_{16}$ = 16-item Quick Inventory of Depressive Symptomatology–Self-Report, VQIDS-SR$_5$ = 5-item Very Quick Inventory of Depressive Symptomatology–Self-Report.

greatest severity of depression for endorsement, while for the VQIDS-SR$_5$, the psychomotor retardation item required the greatest severity of depression. Figure 1 shows that PHQ-A followed by QIDS-SR$_{16}$ and VQIDS-SR$_5$ provide the most precise estimates of symptom severity within 2 standard deviations of the average. Table 4 shows the pairwise conversions of total scores between the VQIDS-SR$_5$ and QIDS-SR$_{16}$, VQIDS-SR$_5$ and PHQ-A, and QIDS-SR$_{16}$ and PHQ-A, respectively.

Are the Scales Sensitive to Change?

Supplementary Figure 2 reveals a strong relationship between QIDS-SR$_{16}$ total score and VQIDS-SR$_5$ ($r = 0.87$, n = 682, $P < .0001$) in assessing changes in severity between baseline and 1-month follow-up. A modest but significant association was found between baseline to 1-month changes in PHQ-A total score and VQIDS-SR$_5$ ($r = 0.50$, n = 682, $P < .0001$) and between PHQ-A and QIDS-SR$_{16}$ total scores ($r = 0.56$; n = 682, $P < .0001$). Effect sizes for the changes in total scores were 0.45, 0.47, and 0.35 for PHQ-A, QIDS-SR$_{16}$, and VQIDS-SR$_5$, respectively. Note that most patients had been in ongoing treatment and would not be expected to change during the 1-month observation period.

DISCUSSION

This report assessed the psychometric properties of the PHQ-A, the QIDS-SR$_{16}$ and the VQIDS-SR$_5$ in a large sample of 12- to 20-year-olds who screened positive for depression and/or suicidal ideation or were in treatment for depression. All 3 scales were unidimensional with high internal consistencies (Cronbach alphas from 0.76–0.86). Similar levels of internal consistency were seen between the PHQ-9 and QIDS-SR$_{16}$ in an adult primary care setting in Singapore[41] (Cronbach α: 0.87 and 0.79, respectively). Also, in an adult sample of 297 depressed inpatients in China, Cronbach α was 0.88 (PHQ-9) and 0.83 (QIDS-SR$_{16}$).[42] The corrected item-total correlations among the 3 instruments were also acceptable: 0.49–0.72 (PHQ-A), 0.29–0.64 (QIDS-SR$_{16}$), and 0.43–0.61 (VQIDS-SR$_5$). Test information function was highest for PHQ-A and acceptable but lower for the VQIDS-SR$_5$, as might be expected with a shorter measure.[43] All 3 scales were sensitive to change over a 1-month follow-up period.

Thresholds for severity categories were nearly identical for PHQ-A and QIDS-SR$_{16}$ (ie, 0–5; 5–10; 10–15; 15–20; 20+) as might be expected since each rating assesses the same 9 criterion symptom domains with each rated 0–3 (range of total score being 0–27 on both ratings). When total score changes between baseline and 1-month follow-up were computed for each scale, the correlation between changes in the VQIDS-SR$_5$ and QIDS-SR$_{16}$ was, as expected, high since the VQIDS-SR$_5$ items are included in the QIDS-SR$_{16}$ total. This finding suggests that VQIDS-SR$_5$ total score is a reasonable proxy for QIDS-SR$_{16}$, albeit with less test information. It cannot be used to screen for all 9 criterion symptom domains as can the PHQ-A or the QIDS-SR$_{16}$. However, a brief assessment may be sufficient when frequent assessment of severity is necessary.

Correlations between changes over the 1-month period in PHQ-A with QIDS-SR$_{16}$ and VQIDS-SR$_5$ were modest (range, 0.50–0.56), suggesting that the change in the prevalence of symptoms reflected in the PHQ-A total score over time is not as tightly tied to the change in symptom severity despite measuring the same 9 criterion symptom domains as the QIDS-SR$_{16}$. This distinction between severity and prevalence is often seen clinically as some persons have a persistent low level of depressive symptoms while others have episodic exacerbations that are impersistent. The expression of disease severity may include both symptom severity and prevalence/pervasiveness, both of which can change over time. A response could entail a reduction in severity, pervasiveness, or both. Frequent sampling of both

severity and pervasiveness may add precision not available with the assessment of only one of these parameters.

Of note, the item-total correlations for the sleep and appetite/weight items on the QIDS-SR_{16} were modest. These results are consistent with a recent report[44] using the Chinese version of the QIDS-SR_{16} in a sample of adolescents with major depressive episodes or bipolar depressive episodes, where item-total correlations were found to be the lowest for the sleep and appetite/weight items as well. This contrasts with the PHQ-A findings in this report in which the item-total correlations were not lower for the sleep and appetite items. An examination of these individual PHQ-A and QIDS-SR_{16} items suggests that the thresholds, which were based on adults, may be less suitable for adolescents. That is, adolescents' regulation of sleep and appetite may be affected by more than depression (stresses with schoolwork, comorbid conditions such as anxiety, etc). It may be easier to estimate them more accurately, and in ways that are consistent with the overall concept of depression, if the measurement period is over 2 weeks, or if prevalence thresholds are used in place of severity thresholds.

Overall, these findings suggest that any of the 3 measures constitute a satisfactory outcome tool. The VQIDS-SR_5 total score is a time-saving alternative that is highly correlated with the QIDS-SR_{16} total score. The present crosswalk tables provide reassurance that one can validly convert QIDS-SR_{16} total to VQIDS-SR_5 total and vice versa. The thresholds identified in this report with younger patients closely approximate those in a prior report that crosswalked the clinician rated VQIDS-C_5 and clinician completed QIDS-C_{16}.[25] Specifically, none, mild, moderate, severe, and very severe category thresholds were 0–1, 2–4, 5–8, 9–12, and 13–15, respectively, with the VQIDS-SR_5, while they were 0–2, 3–5, 6–8, 9–12, and 13–15 with the clinician version of the VQIDS-C_5. Item-total correlations ranged from 0.43 to 0.61 in the youth in this report, while they were 0.57–0.74 with adults using the clinician VQIDS-SR_5. This less consistent performance in this report, though certainly satisfactory, could be due to the shift from clinician to patient rating or the younger age of this sample.

Limitations

This report has several limitations. The QIDS-SR_{16} and VQIDS-SR_5 were derived from the adult IDS-SR, while the PHQ-A was adapted to an adolescent population. Whether a tighter relationship between these measures would have resulted had an adolescent version of the QIDS-SR been used is unknown.[45] The subjects in this report may have already been receiving treatment for depression at the time of the study, so the modest change between the baseline and 1-month assessments may be a result of ongoing care or the fact that the time assessed was limited to 1 month. Longer observation periods with more measurement occasions would be more informative. In addition, many subjects in the study were already receiving pharmacologic treatments that could have affected sleep, appetite/weight, and other symptoms. These medications could have affected both severity and prevalence of symptoms.

There may be a lack of systematic reporting of the comorbidities and, possibly, underdiagnosis of bipolar disorder within this sample. A few participants of the TX-YDSRN study have a diagnosis of bipolar disorder. Being that participants are young and lacking a long-term follow-up (at the time of the collection of this data), fewer episodes can be observed, and this may limit diagnostic accuracy within this sample for bipolar disorders. Furthermore, juvenile onset depressive episodes represent a major risk factor for diagnosis of bipolar disorder.[46] There is also evidence on the importance of assessing subsyndromal mixed/manic features in depressive episodes,[47] which often are correlates of depression severity in adolescence.[48]

It is important to note that the VQIDS-SR_5 does not assess suicidal ideation. Suicidal ideation should be assessed at every visit with all patients receiving psychotropic medications regardless of whether a scale is used or not. Scales administration was not randomized. The PHQ-A was given first, followed by IDS-SR. Completion of one scale could have affected the results of the scale that followed.

CONCLUSION

In summary, PHQ-A, QIDS-SR_{16}, and VQIDS-SR_5 are unidimensional, psychometrically acceptable self-reports that assess depressive symptom burden whether based on severity (QIDS-SR_{16} and VQIDS-SR_5) or on prevalence (PHQ-A) of the 9 criterion symptom domains that define MDD by *DSM-5*. They are sensitive to change over time. Most importantly, total scale scores of any measure can be converted to that of any other measure.

Article Information

Published Online: December 18, 2023. https://doi.org/10.4088/JCP.23m14861

Submitted: March 13, 2023; accepted September 13, 2023.

To Cite: Nandy K, Rush AJ, Carmody T, et al. A comparison of depressive symptom self-reported measures in the Texas Youth Depression and Suicide Research Network (TX-YDSRN). *J Clin Psychiatry.* 2024;85(1):23m14861

Author Affiliations: Peter O'Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, Dallas (Nandy, Carmody); Center for Depression Research and Clinical Care, Peter O'Donnell Jr. Brain Institute and the Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas (Nandy, Carmody, Slater, Mayes, Trivedi); Curbstone Consultant LLC, Santa Fe, New Mexico (Rush); Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, North Carolina (Rush); Duke-National University of Singapore, Singapore (Rush); Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas (Kennard, Emslie); Children's Medical Center, Dallas, Texas (Kennard, Emslie); University of Texas Medical Branch, Galveston (DeFilippis); Department of Psychiatry, University of Texas Rio Grande Valley, School of Medicine, Edinburg (Garza); Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine Houston Texas (Storch); Department of Psychiatry, Texas Tech University Health Science Center, Lubbock, (Wakefield).

Corresponding Author: Madhukar H. Trivedi, MD, Center for Depression Research and Clinical Care, Peter O'Donnell Jr. Brain Institute, Department of Psychiatry, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390-9119 (madhukar.trivedi@utsouthwestern.edu).

Relevant Financial Relationships: Dr 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 Symptomatology and its derivatives). He is named co-inventor on two patents: U.S. 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 U.S. 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. **Dr Carmody** has served as a consultant for Alkermes, Inc. **Dr Kennard** receives royalties from Guilford Press, Inc. and serves on the board of trustees for the Jerry M. Lewis, III, MD Research Foundation. **Dr Emslie** is a consultant for Lundbeck, Neuronetics, and Otsuka and receives research support from American Foundation for Suicide Prevention, Janssen Research & Development and Pharmaceuticals, National Institutes of Health (NIH), Patient-Centered Outcomes Research Institute, and the State of Texas. **Dr Storch** receives grant support from NIH, the Ream Foundation, Greater Houston Community Foundation, International OCD Foundation, and Texas Higher Education Coordinating Board; receives book royalties from Elsevier, Springer, American Psychological Association, Jessica Kingsley, Oxford, and Lawrence Erlbaum; was a consultant for Brainsway and Biohaven; and owns stock in NView. **Dr Trivedi** has provided consulting services to Alkermes Inc, Axsome Therapeutics, Biogen MA Inc., Cerebral Inc., Circular Genomics Inc, Compass Pathfinder Limited, GH Research Limited, Heading Health Inc, Janssen, Legion Health Inc, Merck Sharp & Dohme Corp., Mind Medicine (MindMed) Inc, Merck Sharp & Dhome LLC, Naki Health, Ltd., Neurocrine Biosciences Inc, Noema Pharma AG, Orexo US Inc, Otsuka American Pharmaceutical Inc, Otsuka Canada Pharmaceutical Inc, Otsuka Pharmaceutical Development & Commercialization Inc, Praxis Precision Medicines Inc, SAGE Therapeutics, Sparian Biosciences Inc, Takeda Pharmaceutical Company Ltd, and WebMD; is on the Scientific Advisory Board of Alto Neuroscience Inc, Cerebral Inc., Compass Pathfinder Limited, Heading Health, GreenLight VitalSign6 Inc, Legion Health Inc, Merck Sharp & Dohme Corp, Orexo US Inc, and Signant Health; holds stock in Alto Neuroscience Inc, Cerebral Inc, Circular Genomics Inc, GreenLight VitalSign6 Inc, and Legion Health Inc; and has received editorial compensation from American Psychiatric Association and Oxford University Press. **Drs Nandy**, **Slater**, **Mayes**, **DeFilippis**, **Garza,** and **Wakefield** do not have any conflicts to declare.

Funding/Support: This manuscript was funded by the Texas Youth Depression and Suicide Research Network (TX-YDSRN), a research initiative of the Texas Child Mental Health Care Consortium (TCMHCC). The TCMHCC was created by the 86th Texas Legislature and, in part, funds multi-institutional research to improve mental health care for children and adolescents in Texas. The content is solely the responsibility of the authors and does not necessarily represent the official views of the various funding organizations.

Role of the Funders/Sponsors: The funding source had no role in the design, collection, analysis, or interpretation of the data.

Disclaimer: The Intellectual Property of TX-YDSRN belongs to the University of Texas Southwestern Medical Center (Principal Investigator, Dr Trivedi).

Acknowledgments: The authors thank the patients, clinics, staff, and colleagues who made this project possible. They acknowledge the TX-YDSRN teams from the following sites: University of Texas Southwestern Medical Center, Baylor Medical Center, Texas A&M University System Health Science Center, Texas Tech University Health Science Center Lubbock, Texas Tech University Health Science Center El Paso, University of Texas at Austin Dell Medical School, University of Texas Health San Antonio, University of Texas Rio Grande Valley, University of Texas Health Science Center Houston, University of Texas Health Science Center Tyler, University of Texas Medical Branch, and University of North Texas Health Science Center.

Licensing of QIDS-SR_{16} and VQIDS-SR_5: Licensing and distribution of QIDS-SR_{16} and VQIDS-SR_5 are managed by Mapi Research Trust on behalf of the copyright holder, University of Texas Southwestern Medical Center. Requests for information and licensing of QIDS-SR_{16} and VQIDS-SR_5 should be submitted through Mapi Research Trust's ePROVIDE platform (https://eprovide.mapi-trust.org/).

Supplementary Material: Available at Psychiatrist.com.

References

1. Ghandour RM, Sherman LJ, Vladutiu CJ, et al. Prevalence and treatment of depression, anxiety, and conduct problems in US children. *J Pediatr*. 2019;206:256–267.e3.
2. Ryan ND. Treatment of depression in children and adolescents. *Lancet*. 2005;366(9489):933–940.
3. Thapar A, Collishaw S, Pine DS, et al. Depression in adolescence. *Lancet*. 2012;379(9820):1056–1067. www.ncbi.nlm.nih.gov/pmc/articles/PMC3488279/
4. Thapar A, Eyre O, Patel V, et al. Depression in young people. *Lancet*. 2022;400(10352):617–631.
5. Costello EJ, Maughan B. Annual research review: optimal outcomes of child and adolescent mental illness. *J Child Psychol Psychiatry*. 2015;56(3):324–341.
6. Walter HJ, Abright AR, Bukstein OG, et al. Clinical practice guideline for the assessment and treatment of children and adolescents with major and persistent depressive disorders. *J Am Acad Child Adolesc Psychiatry*. 2022;62(5):479–502.
7. 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. (published correction appears in *JAMA*. 2019 Nov 5;322(17) *JAMA*. 2008;299(8):901–913.
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.
9. Dwyer JB, Stringaris A, Brent DA, et al. Annual research review: defining and treating pediatric treatment-resistant depression. *J Child Psychol Psychiatry*. 2020;61(3):312–332.
10. Poznanski EO, Cook SC, Carroll BJ. A depression rating scale for children. *Pediatrics*. 1979;64(4):442–450.
11. 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.
12. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382–389.
13. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23(1):56–62.
14. Deshpande PR, Rajan S, Sudeepthi BL, et al. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res*. 2011;2(4):137–144.
15. Rush AJ, Giles DE, Schlesser MA, et al. The Inventory for Depressive Symptomatology (IDS): preliminary findings. *Psychiatry Res*. 1986;18(1):65–87.
16. Rush AJ, Gullion CM, Basco MR, et al. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med*. 1996;26(3):477–486.
17. Rush AJ, Rago WV, Crismon ML, et al. Medication treatment for the severely and persistently mentally ill: the Texas Medication Algorithm Project. *J Clin Psychiatry*. 1999;60(5):284–291.
18. 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. (published correction appears in Biol Psychiatry. 2003 Sep 1;54(5) *Biol Psychiatry*. 2003;54(5):573–583.
19. Trivedi MH, Rush AJ, Wisniewski SR, et al; STAR*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry*. 2006;163(1):28–40.
20. Fortney JC, Unützer J, Wrenn G, et al. A tipping point for measurement-based care. *Psychiatr Serv*. 2017;68(2):179–188.
21. Scott K, Lewis CC. Using measurement-based care to enhance any treatment. *Cognit Behav Pract*. 2015;22(1):49–59.
22. Rush AJ. Narrowing the gaps between what we know and what we do in psychiatry. *J Clin Psychiatry*. 2015;76(10):1366–1372.
23. 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(1-3):115–118.
24. Rush AJ, Aaronson ST, Demyttenaere K. Difficult-to-treat depression: a clinical and research roadmap for when remission is elusive. *Aust N Z J Psychiatry*. 2019;53(2):109–118.
25. Rush AJ, Sackeim HA, Conway CR, et al. Clinical research challenges posed by difficult-to-treat depression. *Psychol Med*. 2022;52(3):419–432.
26. De La Garza N, John Rush A, Grannemann BD, et al. Toward a very brief self-report to assess the core symptoms of depression (VQIDS-SR_5). *Acta Psychiatr Scand*. 2017;135(6):548–553.
27. Rush AJ, Madia ND, Carmody T, et al. Psychometric and Clinical Evaluation of the Clinician (VQIDS-C_5) and Self-Report (VQIDS-SR_5) Versions of the Very Quick Inventory of Depressive Symptoms. *Neuropsychiatr Dis Treat*. 2022;18:289–302.
28. Trivedi MH, Minhajuddin A, Slater H, et al. Texas Youth Depression and Suicide Research Network (TX-YDSRN) research registry and learning healthcare network: rationale, design, and baseline characteristics. *J Affect Disord*. 2023;340:88–99.
29. Johnson JG, Harris ES, Spitzer RL, et al. 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.
30. Wakefield JC. *DSM-5*: an overview of changes and controversies. *Clin Soc Work J*. 2013;41(2):139–154.
31. Bech P, Gram LF, Dein E, et al. Quantitative rating of depressive states. *Acta Psychiatr Scand*. 1975;51(3):161–170.
32. Bech P, Allerup P, Gram LF, et al. The Hamilton depression scale: evaluation of objectivity using logistic models. *Acta Psychiatr Scand*. 1981;63(3):290–299.
33. Bech P, Fava M, Trivedi MH, et al. Factor structure and dimensionality of the two depression scales in STAR*D using level 1 datasets. *J Affect Disord*. 2011;132(3):396–400.

34. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for *DSM-IV* and ICD-10. *J Clin Psychiatry*. 1998;59(suppl 20):22–33, quiz 34–57.
35. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. New York, NY: McGraw Hill; 1994.
36. Humphreys LG, Montanelli RG Jr. An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behav Res*. 1975;10(2):193–205.
37. Samejima F. Graded Response Model. In: Van Der Linden WJ, Hambleton RK, eds. *Handbook of Modern Item Response Theory*. Springer; 1997:85–100.
38. Thissen D, Chen WH, Bock RD. *MULTILOG 7 for Windows: Multiple Category Item Analysis and Test Scoring using Item Response Theory [Computer Program]*. Lincolnwood, IL: Scientific Software International; 2003.
39. Orlando M, Sherbourne CD, Thissen D. Summed-score linking using item response theory: application to depression measurement. *Psychol Assess*. 2000;12(3):354–359.
40. Birnbaum A. Some latent trait models and their use in inferring an examinee's ability. In: Lord FM, Novick MR, eds. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley; 1968.
41. Sung SC, Low CC, Fung DS, et al. Screening for major and minor depression in a multiethnic sample of Asian primary care patients: a comparison of the nine-item Patient Health Questionnaire (PHQ-9) and the 16-item Quick Inventory of Depressive Symptomatology - Self-Report (QIDS-SR16). *Asia-Pac Psychiatry*. 2013;5(4):249–258.
42. Feng Y, Huang W, Tian TF, et al. The psychometric properties of the Quick Inventory of Depressive Symptomatology-Self-Report (QIDS-SR) and the Patient Health Questionnaire-9 (PHQ-9) in depressed inpatients in China. *Psychiatry Res*. 2016;243:92–96.
43. Baker F. *The Basics of Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation*. University of Maryland; 2001.
44. 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.
45. Haley CL, Kennard BD, Bernstein IH, et al. Improving Depressive Symptom Measurement in Adolescents: A Psychometric Evaluation of the Quick Inventory of Depressive Symptomatology, Adolescent Version (Qids-A17); 2009, unpublished dissertation.
46. Faedda GL, Marangoni C, Serra G, et al. Precursors of bipolar disorders: a systematic literature review of prospective studies. *J Clin Psychiatry*. 2015;76(5):614–624.
47. Diler RS, Goldstein TR, Hafeman D, et al. Distinguishing bipolar depression from unipolar depression in youth: preliminary findings. *J Child Adolesc Psychopharmacol*. 2017;27(4):310–319.
48. Serra G, Iannoni ME, Trasolini M, et al. Characteristics associated with depression severity in 270 juveniles in a major depressive episode. *Brain Sci*. 2021;11(4):440.

The Journal of Clinical Psychiatry

Supplementary Material

Article Title: A Comparison of Depressive Symptom Self-Reported Measures in the Texas Youth Depression and Suicide Research Network (TX-YDSRN)

Authors: Karabi Nandy, PhD; A. John Rush, MD; Thomas Carmody, PhD; Beth D. Kennard, MD; Graham J. Emslie, MD; Holli Slater, PhD; Taryn L. Mayes, MS; Melissa DeFilippis, MD; Cynthia Garza, PsyD; Eric A. Storch, PhD; Sarah M. Wakefield, MD; and Madhukar H. Trivedi, MD

DOI Number: 10.4088/JCP.23m14861

LIST OF SUPPLEMENTARY MATERIAL FOR THE ARTICLE

DISCLAIMER

This Supplementary Material has been provided by the author(s) as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

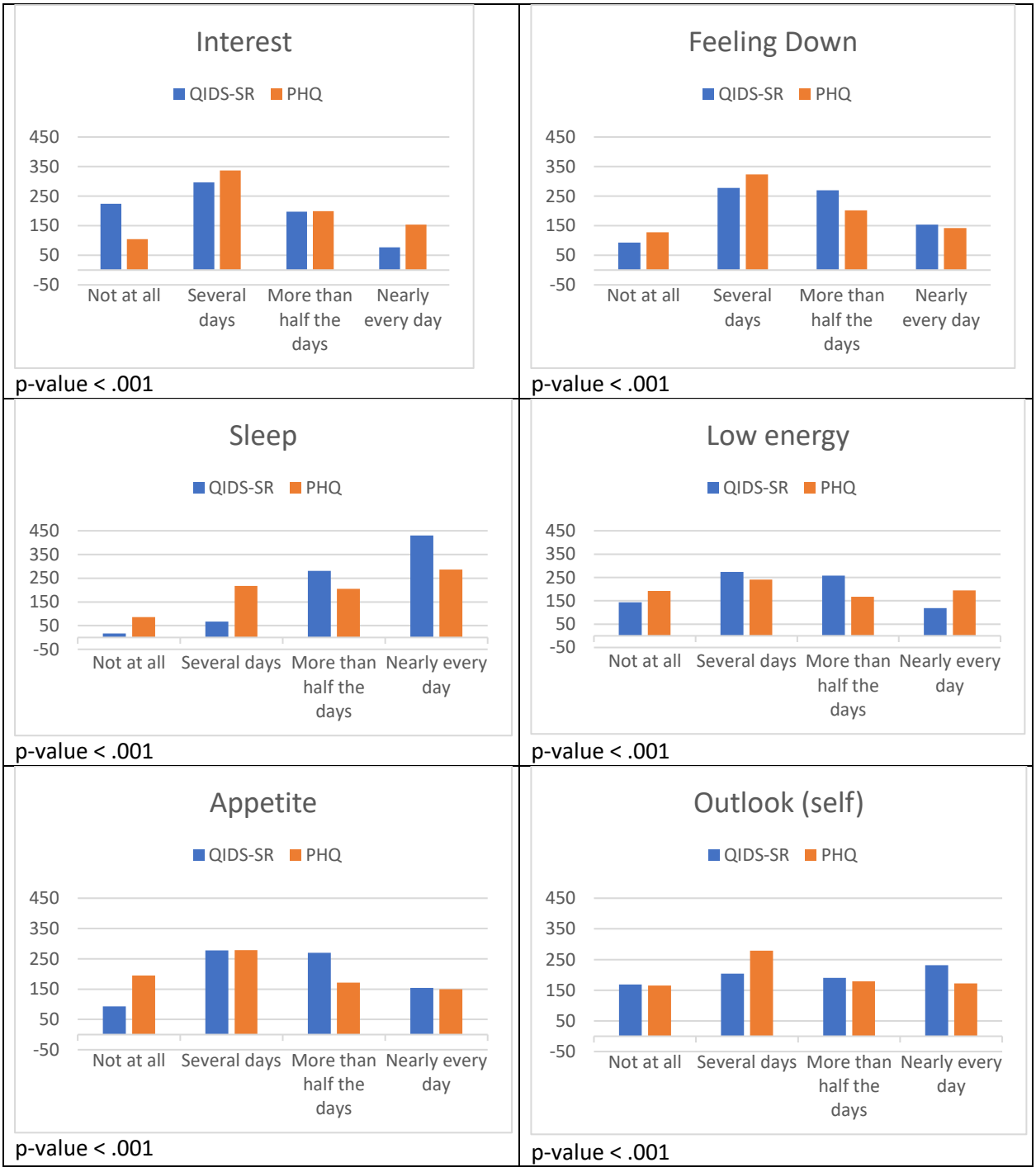
SUPPLEMENTARY TABLES AND FIGURES

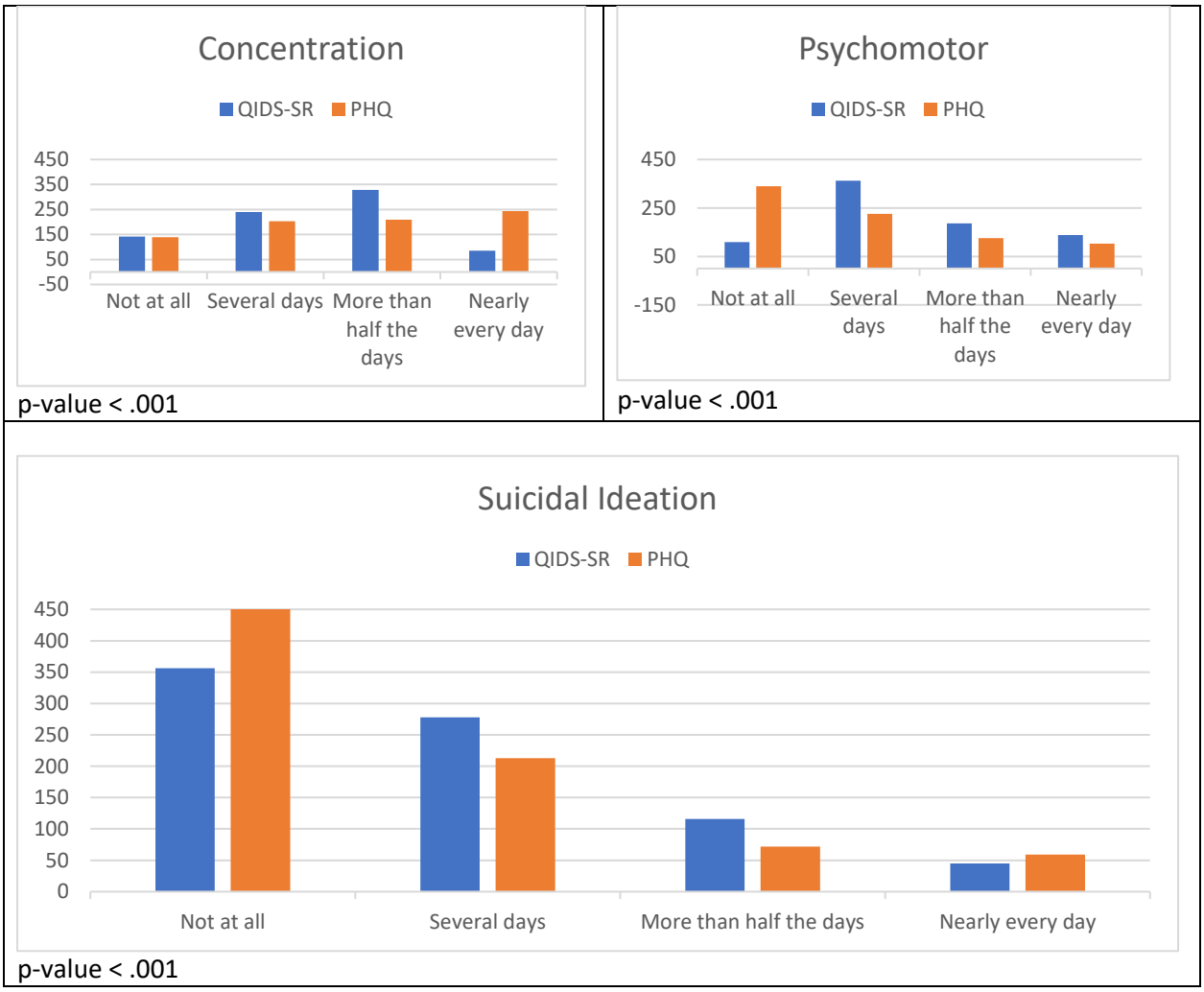
Supplementary Table 1. List of items in IDS-SR, QIDS-SR$_{16}$, PHQ-A, VQIDS-SR$_5$ (all items are scores 0-3)

30-item IDS-SR	16-item QIDS-SR$_{16}$	PHQ-A (matching 9 domains of PHQ-A to the 9 domains in QIDS-SR$_{16}$)	5-item VQIDS-SR$_5$
1. Sleep onset insomnia	√	Item 3: Trouble falling or staying asleep, or sleeping too much	
2. Mid-nocturnal insomnia	√	Item 3	
3. Early morning insomnia	√	Item 3	
4. Hypersomnia	√	Item 3	
5. Mood (sad)	√	Item 2: Feeling down, depressed, or hopeless	√
6. Mood (irritable)			
7. Feeling anxious or tense			
8. Response of Mood to Good or Desired Events			
9. Mood in relation to time of day			
10. Quality of mood			
11. Appetite (decreased)	√	Item 5: Poor appetite or overeating	
12. Appetite (increased)	√	Item 5	
13. Weight (decrease) within the last two weeks	√	Item 5	
14. Weight (increase) within the last two weeks	√	Item 5	
15. Concentration/decision making	√	Item 7: Trouble concentrating	
16. Outlook (self)	√	Item 6: Feeling bad about yourself – or that you are a failure	√
17. View of future			
18. Suicidal ideation (Thoughts of Death or Suicide)	√	Item 9: Thoughts that you would be better off dead or of hurting yourself	
19. Involvement (General Interest)	√	Item 1: Little interest or loss of pleasure	√
20. Energy/fatigability	√	Item 4: Feeling tired or having little energy	√
21. Capacity for Pleasure or Enjoyment (excluding sex)			
22. Interest in Sex			
23. Psychomotor slowing (Feeling slowed down)	√	Item 8: Moving or speaking so slowly that other people could notice	√
24. Psychomotor agitation (Feeling restless)	√	Item 8	
25. Aches and pains			
26. Other bodily symptoms			
27. Panic/Phobic symptoms			
28. Constipation/diarrhea			
29. Interpersonal Sensitivity			
30. Leaden Paralysis/Physical Energy			
IDS total score = sum of items (5-10, 15-22, 25 + 26-30) + Q1 + Q2 + Q3; where Q1=Max of 4 sleep items (1 to 4) Q2=Max of 4 appetite/weight items (11 to 14)	QIDS total score = item 5 + item 15 + item 16 + item 18 + item 19 + item 20 + Q1 + Q2 + Q3; where Q1=Max of 4 sleep items (1 to 4) Q2=Max of 4 appetite/weight items (11 to 14)	PHQ Total Score = sum of items 1-9	VQIDS total score = item 5 + item 16 + item 19 + item 20 + item 23

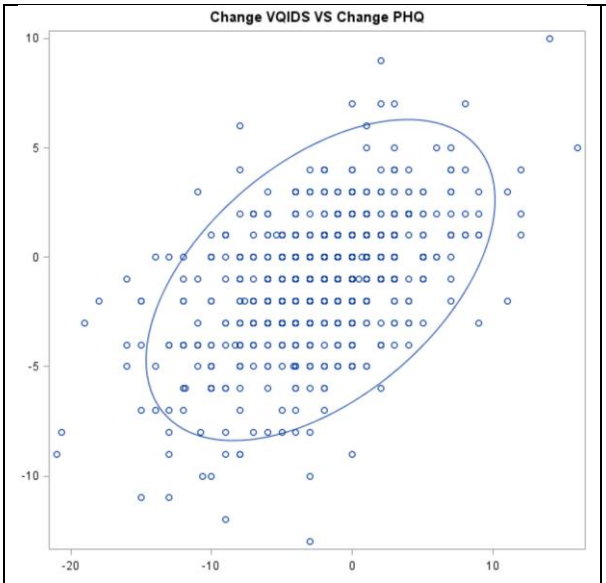
Q3=Max of 2 psychomotor items (23 and 24)	Q3=Max of psychomotor items (23 and 24)		

Supplementary Figure 1. Frequencies of QIDS-SR$_{16}$ and PHQ-A Ratings by Domain (n=795)

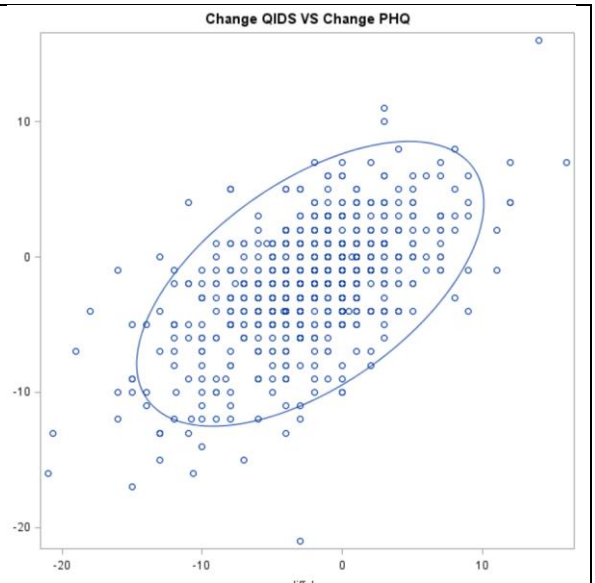




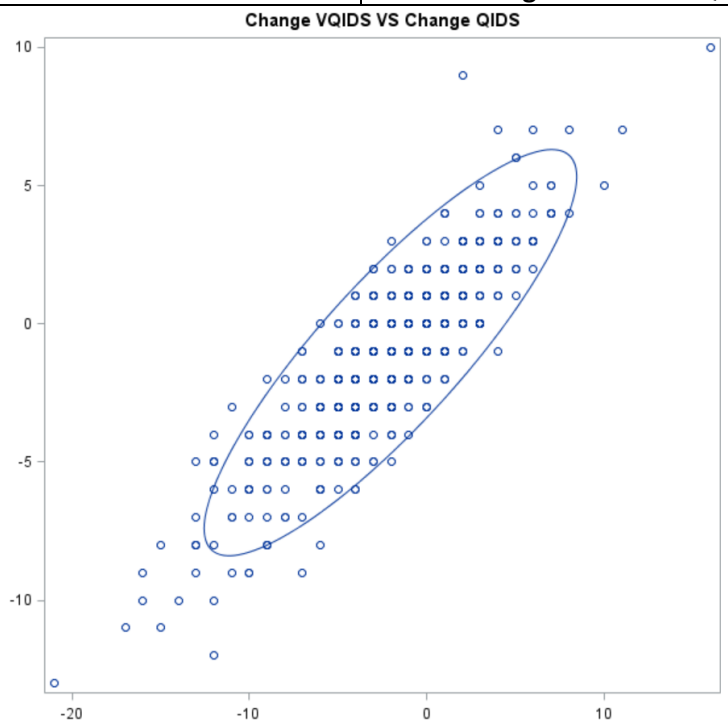
Supplementary Figure 2: Scatterplots of Change Scores of Pairs of PHQ-A, QIDS-SR$_{16}$ and VQIDS-SR$_5$ total scores (n=682)



Change score of PHQ-A on x-axis



Change score of PHQ-A on x-axis



Change score of QIDS-SR_{16} on x-axis