HHS Public Access

Author manuscript

Behav Ther. Author manuscript; available in PMC 2020 March 01.

Published in final edited form as:

Behav Ther. 2019 March ; 50(2): 314–324. doi:10.1016/j.beth.2018.06.003.

Defining Treatment Outcomes in Pediatric Obsessive-Compulsive Disorder Using a Self-report Scale

Joseph F. McGuire, Ph.D.[a], **Daniel A Geller, M.D.**[b], **Tanya K. Murphy, M.D.**[c], **Brent J. Small, Ph.D.**[d], **Arianna Unger**[b], **Sabine Wilhelm, Ph.D.**[b], and **Eric A. Storch, Ph.D.**[e]

[a]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine

[b]Massachusetts General Hospital and Harvard Medical School

[c]Departments of Pediatrics, Psychiatry and Behavioral Neurosciences, University of South Florida

[d]School of Aging Studies, University of South Florida

[e]Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine

Abstract

This study examined benchmarks of treatment response and clinical remission on the Obsessive Compulsive Inventory-Child Version (OCI-CV) for youth with obsessive-compulsive disorder (OCD). Participants were 91 youth who enrolled in a randomized controlled trial that examined the benefit of augmenting cognitive behavior therapy (CBT) with either d-cycloserine or placebo. Youth completed the OCI-CV at baseline, Week 4 (prior to initiating exposure therapy), and post-treatment. Receiver operator curve (ROC) analyses examined optimal benchmarks for treatment response and clinical remission as identified by independent evaluators at the post-treatment assessment using the Clinical Global Impression (CGI) scales of Improvement (CGI-Improvement), Severity (CGI-Severity), and Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS). Optimal benchmarks for treatment response were a 20-25% reduction in the OCI-CV total score. Meanwhile, optimal benchmarks for remission were a 55-65% reduction in the OCI-CV total score and a post-treatment total score ≤6-8. OCI-CV benchmarks exhibited moderate agreement with the CY-BOCS for treatment response and clinical remission. Meanwhile, fair agreement was observed for response and remission with CGI scales. A lower pre-treatment OCI-CV total score was associated with less agreement between classification approaches. Findings provide benchmarks for classifying treatment response and clinical remission in an efficient manner. Given the moderate agreement between the CY-BOCS and OCI-CV benchmarks, the OCI-CV may serve as a useful alternative when clinician-rated scales cannot be administered due to limited resources (e.g., time, training). Thus, evidence-based measurement can be incorporated to monitor therapeutic response and remission in clinical practice.

Corresponding Author: Joseph F. McGuire, Ph.D., Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, Maryland. 1800 Orleans St, Baltimore, MD 21287. jfmcguire@jhmi.edu.

Keywords

children; adolescents; obsessive-compulsive disorder; cognitive-behavior therapy; evidence-based assessment

Introduction

Obsessive-compulsive disorder (OCD) predominantly emerges during childhood (Ruscio, Stein, Chiu, & Kessler, 2010), causes impairment (Piacentini, Bergman, Keller, & McCracken, 2003), and is associated with a poor quality of life (Storch et al., 2017). Evidence-based treatments for youth with OCD include exposure-based cognitive behavior therapy (CBT) and pharmacotherapy with serotonin reuptake inhibitors (SRIs) (McGuire et al., 2015). Experts recommend CBT as the first line treatment for youth with mild-to-moderate OCD, and in combination with SRI medications for youth with moderate-to-severe OCD symptoms (Geller, March, & the AACAP Committee on Quality Issues, 2012). Thus, in practically every case of pediatric OCD, it is recommended that youth receive exposure-based CBT.

While randomized clinical trials (RCTs) demonstrate the efficacy of CBT and SRIs, applying RCT methodology to monitoring treatment outcomes in clinical practice is challenging. In RCTs of pediatric OCD, treatment response is determined by an independent evaluator naïve to treatment condition using the Clinical Global Impression of Improvement (CGI-Improvement; Guy, 1976), and clinical remission classified by the discontinuation of diagnostic criteria on a diagnostic interview (Kaufman et al., 1997). While relevant in clinical trials research, incorporating these outcome measures into clinical practice proves difficult for multiple reasons (e.g., time demands on provider and family, required training). Fortunately, efforts have been made to benchmark treatment response and clinical remission using clinician-rated measures of pediatric OCD such as the Children's Yale-Brown Obsessive-Compulsive Scale (CY-BOCS; Scahill et al. 1997).

The CY-BOCS is a clinician-rated measure of OCD symptom severity, which consists of a symptom checklist that assesses common obsessions and compulsions (Scahill et al., 1997). Thereafter, the clinician inquires about current OCD symptoms to complete a 10-item severity scale, which integrates this information to quantify symptom severity guided by anchor points. This measure is typically administered to both youth and parents together. Research suggests that a 25-35% reduction on the CY-BOCS total score is associated with a positive treatment response, with a 45-55% reduction on the CY-BOCS total score (or total score cutoff ≤11-14) associated with clinical remission (Skarphedinsson, De Nadai, Storch, Lewin, & Ivarsson, 2017; Storch, Lewin, De Nadai, & Murphy, 2010). Although offering quantitative benchmarks for categorizing treatment response and clinical remission, administration of the CY-BOCS in clinical practice is challenging due to the time-intensive nature of the measure (~30 minutes), required training for administration, and non-linear nature of the scale. Therefore, there remains a continued need for accurate, efficient, and effective measures of pediatric OCD that can be easily implemented in clinical practice.

In comparison to diagnostic interviews and clinician-administered scales, self-report rating scales offer several advantages in identifying OCD symptoms and quantifying severity in an efficient manner. These measures are brief, cost-effective, require minimal training to administer and interpret, and have the advantage of removing potential clinician bias. Several child-report rating scales exist for pediatric OCD (see Rapp, Bergman, Piacentini, & McGuire, 2016 for a review). A widely-used measure in research and clinical practice is the Obsessive Compulsive Inventory-Child Version (OCI-CV; Foa et al., 2010). The OCI-CV is a 21-item questionnaire that assesses the presence of obsessive-compulsive symptoms over the past month, and yields a total score that has good psychometric properties (Foa, et al., 2010; Jones et al., 2013). However, there has been no formal evaluation of benchmarks associated with treatment response and clinical remission on the OCI-CV. This information is essential to assist clinicians in determining when youth have achieved a clinically meaningful improvement in an accurate and efficient manner, which is critical for treatment planning.

Given the importance of incorporating evidence-based measurement into evidence-based treatment (Ng & Weisz, 2016), we sought to characterize benchmarks of treatment response and clinical remission on the OCI-CV using data from a large-scale multi-site RCT of pediatric OCD. First, we investigated the optimal percent reduction in OCI-CV total score for predicting treatment response. Second, we examined the optimal percent reduction in total score and total score cutoff on the OCI-CV for predicting clinical remission. Third, we examined agreement in classifying treatment response and clinical remission using benchmarks on the OCI-CV, CGI scales, and CY-BOCS. Finally, we explored demographic and clinical characteristics that influenced agreement between classification methods.

Methods

Participants

In this multi-site RCT, 142 youth with OCD were recruited from two sites (Massachusetts General Hospital and the University of South Florida) and randomly assigned to receive exposure-based CBT augmented with either d-cycloserine or placebo (Storch et al., 2016). To participate, youth must have had a confirmed diagnosis of OCD, a CY-BOCS total score ≥16, and a full scale IQ ≥85. Concurrent psychiatric medications were permissible provided that the medications were stable for up to eight weeks prior to treatment (only six weeks for antipsychotics) and remained stable throughout treatment. Full inclusion and exclusion criteria can be reviewed elsewhere (McGuire et al., 2012). Fifty one youth (36%) were missing ≥15% of OCI-CV items at either baseline, Week 4 (prior to initiating exposure therapy), or post-treatment. Youth who had ≤15% of OCI-CV items missing at all three assessments were included. Although included youth were slightly older than excluded youth (M = 11.57, SD = 2.76 vs. M = 12.89, SD = 2.97, t = 2.61, p < .01), there were no other differences in demographic characteristics (p = .53 - .92) or baseline CY-BOCS total score (p = .89) between these groups.

The final sample of 91 youth was approximately 13 years of age (M = 12.89, SD = 2.97, range: 7-17), predominantly Caucasian (87.9%), and had a near equivalent gender distribution (53.8% female). These youth had moderate OCD symptom severity at baseline

on the CY-BOCS total score (M = 25.25, SD = 5.30), and 30.8% were taking selective serotonin reuptake inhibitors that were stable over the study. Common co-occurring psychiatric conditions included: anxiety disorders (e.g., separation anxiety, social phobia, generalized anxiety disorder; 49.5%), attention deficit/hyperactivity disorder (ADHD; 25.3%), major depressive disorder (16.5%), and persistent tic disorders (e.g., persistent tic disorder or Tourette Disorder, 7.7%).

Measures

Obsessive Compulsive Inventory-Child Version (OCI-CV; Foa, et al., 2010). The OCI-CV is a 21-item child-report scale that assesses the presence and frequency of obsessive-compulsive symptoms over the past month. Items are rated on a 3-point scale and summed to produce a total score (range: 0-42). The OCI-CV total score has good psychometric properties (Foa, et al., 2010; Jones, et al., 2013; Rodriguez-Jimenez et al., 2016), and treatment sensitivity (Foa, et al., 2010). The internal consistency of the OCI-CV total score was good across all assessments (α = .86 - .88), with good test-retest reliability observed between baseline and the Week 4 assessment (intraclass correlation = .85). There was small non-significant associations between the OCI-CV total score and the CY-BOCS total score (r = .06, p = .55) and CGI-Severity (r = .17, p = .10), which improved by post-treatment (CY-BOCS total score, r = .43, p < .001; CGI-Severity, r = .40, p < .001).

Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS; Scahill, et al., 1997). The CY-BOCS is a clinician-administered interview that assesses OCD symptom severity over the past week. The CY-BOCS total score has demonstrated good reliability, validity, and treatment sensitivity (Scahill, et al., 1997; Skarphedinsson, et al., 2017; Storch, et al., 2010). For this evaluation, conservative benchmarks were used to categorize treatment response (≥35% reduction in total score) and clinical remission (≥55% reduction in total score, or total score cutoff ≤11; Skarphedinsson, et al., 2017).

Clinical Global Impression of Improvement (CGI-Improvement) and Severity (CGI-Severity) Scales (Guy, 1976). The CGI-Improvement is a clinician-rated scale that assesses global improvement from baseline, which ranges from "very much improved" (1) to "very much worse" (7). A score of "much improved" or "very much improved" is considered a positive treatment response in pediatric OCD RCTs. The CGI-Severity is clinician-rated scale that provides a global rating of OCD severity, which ranges from "not at all ill" (0) to "extremely ill" (6). Consistent with similar evaluations (Skarphedinsson, et al., 2017; Storch, et al., 2010), scores of "no illness" or "slight illness" were used to designate clinical remission.

Procedures

All study procedures were approved by the institutional review boards at the two recruitment sites. After explaining study procedures, written parental consent and youth assent were obtained. At the screening assessment, youth and parents were administered the CY-BOCS and a structured diagnostic interview to ascertain study eligibility. Eligible and interested youth subsequently completed a baseline assessment that included the CY-BOCS administered to parents and youth, and youth completed the OCI-CV. After completing the

baseline assessment, youth received 3 CBT sessions focused on psychoeducation, hierarchy development, and cognitive restructuring of obsessive-compulsive thoughts. Prior to the 4th session in which exposure and response prevention therapy started (Week 4), youth were re-assessed on the CY-BOCS and were re-administered the OCI-CV. Youth who continued to meet inclusion criteria were randomly assigned to receive 7 weekly sessions of protocol-driven CBT that were paired either with d-cycloserine or placebo. After CBT, youth and parents completed a post-treatment assessment that included the CY-BOCS, CGI-Improvement, CGI-Severity, and the OCI-CV. All clinical evaluators and therapists were blind to treatment condition.

Analytic Plan

Data were missing completely at random (Little's MCAR, p=.37) and constituted three missing items for three different participants. The three missing item responses were addressed using expectation maximization (Peugh & Enders, 2004). This approach provides accurate estimates for replacing missing data when ≤15% of item-level data are missing (Enders, 2010).

Descriptive statistics characterized the sample. The performance of the OCI-CV total score in predicting treatment response with the CGI-Improvement and CY-BOCS, and remission with the CGI-Severity and CY-BOCS was assessed using receiver operating characteristic (ROC) statistics. ROC statistics examined the ratios of true-positive, false-positive, true-negative, and false-negative results. Consistent with prior studies (Skarphedinsson, et al., 2017; Storch, et al., 2010), ROC parameters included: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), Youden's J, and efficiency.

Additionally, quality receiver operator characteristic (QROC) methods were used to address concerns regarding the application of ROC statistics in biomedical research (e.g., error in measurement of gold standard criterion;(Kraemer, 1992). QROC statistics provide specific forms of weighted kappa statistics to measure the quality of specificity [$_k$(0.0)], quality of efficiency [$_k$(0.5)], and quality of sensitivity [$_k$(1.0)]. For these statistics, a value of 0 indicates that the classification of response/remission is no different than chance, whereas a value of 1 indicates perfect classification agreement. Consistent with prior OCD studies (Skarphedinsson, et al., 2017; Storch, et al., 2010), quality of efficiency was prioritized for selecting the optimal benchmark scores. This selects the most efficient cutoff that minimizes the false-positives and false-negatives equally. Consistent with similar reports (Johnco, Salloum, Lewin, & Storch, 2015), if multiple benchmarks had the same maximum quality of efficiency [$_k$(0.5)], then specificity was selected as a secondary criterion. The OCI-CV total score percent reductions were divided into 5% intervals, with total score cutoffs selected in increments of change on a single item (two points). Chi-square and kappa statistics examined agreement between responder/remitter classifications using the OCI-CV, CGI-Improvement, CGI-Severity, and CY-BOCS. Classification agreement using the kappa statistic was categorized using criteria put forth by Landis & Koch (1977): values ≤ 0 as indicating no agreement, 0.01–0.20 as none to slight, 0.21–0.40 as fair, 0.41– 0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement. Chi-square and t-tests explored differences in demographic (e.g., age, gender) and clinical

characteristics (e.g., comorbidity, OCI-CV total score, CY-BOCS total score) between youth who were classified the same or differently between approaches.

Results

Baseline Characteristics and Treatment Outcomes

There were no differences in demographics or OCI-CV scores between recruitment sites at baseline (p = .14 - .98). Eight-one participants (90%) endorsed symptoms on three or more symptom categories on the OCI-CV at the baseline assessment. The OCI-CV subscale scores included: Doubting/Checking (M = 3.99, SD = 3.06, range: 0-10), Obsessing (M = 3.77, SD = 2.61, range; 0-8), Hoarding (M = 1.64, SD = 1.74, range: 0-6), Washing (M = 2.68, SD = 2.27, range: 0-6), Ordering (M = 2.40, SD = 1.98, range: 0-6), and Neutralizing (M = 1.67, SD = 1.67, range: 0-6). The average OCI-CV total score at baseline was 16 (M = 16.14, SD = 8.16, range: 0-36) and post-treatment was 10 (M = 10.21, SD = 7.42, range: 0-29). There was a significant reduction on the OCI-CV total score from pre-treatment to post-treatment, $t(90) = 6.84$, $p < .001$, d = .72, with an average OCI-CV total score percent reduction of 31%.

Seventy-seven youth exhibited a treatment response on the CGI-Improvement (85%), but only 19 youth (21%) exhibited clinical remission on the CGI-Severity. From pre-treatment to post-treatment, there was a significant reduction in CY-BOCS total scores, $t(90) = 17.09$, $p < .001$, d = 1.79. Using conservative benchmarks on the CY-BOCS, 62 youth (68%) exhibited a treatment response and 35-39 youth (39-43%) experienced clinical remission using the percent reduction and total score cutoff, respectively. There was no difference in therapeutic improvement on the CY-BOCS total score [$t(90) = 1.26$ p = .21] or treatment response on the CGI-Improvement ($\chi^2 = 0.39$, p = .53) between CBT augmented with d-cycloserine and placebo. Finally, there was no difference in therapeutic improvement on the OCI-CV total score from pretreatment (DCS: M = 16.39, SD = 7.84; PBO: M = 15.88, SD = 8.56) and post-treatment (DCS: M = 11.65, SD = 8.07; PBO: M = 8.73, SD = 6.44), $F(1, 89) = 1.95$, p= 17.

Predicting Treatment Response

Table 1 presents the ROC/QROC statistics used to predict treatment response. Compared to the CGI-Improvement, maximum quality of efficiency was found at ≥25% reduction. This classified 57 youth (63%) as treatment responders, and produced a false-positive rate of 29% and false-negative rate of 31%. Meanwhile, compared to a ≥35% reduction in CY-BOCS total score, maximum quality of efficiency was found at a ≥20% reduction in OCI-CV total score. This classified 61 youth (67%) as treatment responders, and produced a false-positive rate of 38% and false-negative rate of 19%.

Predicting Clinical Remission

Table 2 presents the ROC/QROC statistics used to predict clinical remission. Compared to the CGI-Severity, maximum quality of efficiency was found at ≥55% reduction in OCI-CV total score. This classified 29 youth (32%) as experiencing clinical remission, and produced a false-positive rate of 24% and false-negative rate of 37%. Meanwhile, compared to a ≥55%

reduction in CY-BOCS total score, maximum quality of efficiency was found at a ≥65% reduction in OCI-CV total score. This classified 23 youth (25%) as experiencing clinical remission, and produced a false-positive rate of 14% and false-negative rate of 57%.

Table 3 presents the ROC/QROC statistics used to predict clinical remission using cutoff scores. Compared to the CGI-Severity, maximum quality of efficiency was found at a total score cutoff of ≥6. This classified 31 youth (34%) as experiencing clinical remission, and produced a false-positive rate of 26% and false-negative rate of 37%. However, compared to a CY-BOCS total score cutoff ≤11, maximum quality of efficiency was found at a total score cutoff of ≤8. This classified 44 youth (48%) as experiencing clinical remission, and produced a false-positive rate of 29% and false-negative rate of 26%.

Agreement and Disagreement Between Treatment Response Classifications

Fair agreement was observed between the OCI-CV ≥25% total score reduction and the CGI-Improvement, which consistently classified 63 youth (69%) as either a responder or nonresponder (see Table 4). Meanwhile, moderate agreement was observed between the OCI-CV ≥20% total score reduction and ≥35% reduction in CY-BOCS total score, which consistently classified 68 youth (75%, see Table 4). Finally, moderate agreement was observed between the ≥35% reduction in CY-BOCS total score and CGI-Improvement, which consistently classified 76 youth (84%, see Table 4).

Twenty-eight youth were classified as a treatment responder by either the OCI-CV ≥25% total score reduction or CGI-Improvement, with most youth classified as a responder on the CGI-Improvement and non-responder on the OCI-CV ($n = 24$). Youth who exhibited disagreement in classification approaches had lower pre-treatment OCI-CV total scores ($M = 13.61$, $SD = 7.73$) compared to youth who exhibited agreement ($M = 17.26$, $SD = 8.16$, $t = 2.00$, $p<.05$, $d = .45$). No other differences in demographic or clinical characteristics were identified between groups ($p = .07$-$.96$). Meanwhile, 23 youth were classified as a treatment responder by either the OCI-CV ≥20% reduction in total score or CY-BOCS. Disagreement was similarly distributed between youth rated as a responder on the CY-BOCS only ($n = 12$) and the OCI-CV only ($n = 11$). There were no differences in demographic or clinical characteristics between youth who exhibited agreement and disagreement in classification approaches ($p = .08$ - 1.00).

Agreement and Disagreement Between Clinical Remission Classifications

Fair agreement was observed between the OCI-CV ≥55% total score reduction and the CGI-Severity, which consistently classified 67 youth (74%) as either a remitter or non-remitter (see Table 4). Similarly, fair agreement was also observed between the OCI-CV total score cutoff of ≤6 and CGI-Severity, which consistently classified 65 youth (72%, see Table 4). Meanwhile, fair agreement was observed between the OCI-CV ≥65% total score reduction and a CY-BOCS ≥55% total score reduction, which consistently classified 63 youth (69%, see Table 4). Comparatively, moderate agreement was found between the OCI-CV total score cutoff ≤8 and CY-BOCS total score cutoff ≤11, which consistently classified 66 youth (73%, see Table 4). Finally, moderate agreement was observed between a CY-BOCS ≥55% reduction in total score and CGI-Severity, which comparably classified 73 youth (80%, see

Table 4). Moderate agreement was also observed between a CY-BOCS total score cutoff ≥11 and CGI-Severity, which consistently classified 69 youth (76%, see Table 4).

Twenty-four youth were classified as experiencing clinical remission by either the OCI-CV ≥55% total score reduction or CGI-Severity, with most youth being rated in remission on the OCI-CV but not the CGI-Severity ($n = 17$). Boys exhibited a higher incidence of classification disagreement compared to girls (38% vs. 16%, $\chi^2 = 5.52$, $p < .02$). No other differences in demographic or clinical characteristics were identified between groups ($p = .15$ - .94). Similarly, 26 youth were classified as experiencing clinical remission on the OCI-CV total score cutoff of ≤6 or the CGI-Severity, with most youth rated as in remission on the OCI-CV but not the CGI-Severity ($n = 19$). No differences in demographic or clinical characteristics were identified between youth who exhibited agreement and disagreement in classification methods ($p = .06$-.99).

Twenty-eight youth were classified as in clinical remission by either the OCI-CV ≤65% total score reduction or CY-BOCS ≤55% total score reduction, with most youth being rated in remission on the CY-BOCS but not the OCI-CV ($n = 20$). Youth who exhibited disagreement in classification approaches had lower pre-treatment OCI-CV total scores ($M = 13.60$, $SD = 8.16$) compared to youth who exhibited agreement ($M = 17.27$, $SD = 7.97$, $t = 2.05$, $p < .05$, $d = .45$). No other differences in demographic or clinical characteristics were identified between groups (p = .12 - 1.00). Finally, 25 youth were classified as in clinical remission by either the OCI-CV total score cutoff ≤6 or CY-BOCS total score cutoff ≤11, with slightly more youth being classified as in remission on the OCI-CV and not the CY-BOCS ($n = 15$). No differences between youth who exhibited agreement and disagreement in classification methods were identified on demographic and clinical characteristics ($p = .07$ - .86).

Discussion

Given the importance of utilizing evidence-based measurement in clinical practice (Ng & Weisz, 2016), this paper examined benchmarks of treatment response and clinical remission using the OCI-CV. Similar to clinician-administered OCD scales benchmarked against the CGI-Improvement and CGI-Severity (Skarphedinsson, et al., 2017; Storch, et al., 2010), a 25% reduction in total score was optimal for determining treatment response, with a 55% reduction optimal for remission. Additionally, an OCI-CV total score ≤6 was found to be optimal for determining clinical remission. Although the sensitivity (.63 - .71), specificity (.71 - .76), PPV (.60 - .63), NPV (.76 - .78), efficiency (.70 - .71) and Youden's J (.37 - .40) were similar across optimal benchmarks, they were somewhat lower than the sensitivity (.82 - .98), specificity (.72 - .90), PPV (.83 - .96), NPV (.72 - .86), and efficiency (.83 - .94) identified in benchmark studies using the CY-BOCS. These differences may be attributed to rater training included in clinician-administered assessments, and/or the utilization of a composite parent and child report in making clinician ratings. Moreover, in prior studies, the same clinician completing the CY-BOCS also completed the CGI ratings. Therefore, it is not surprising that these clinician-rated measures exhibited better predictive agreement with one another in comparison to a child-report scale, as ratings were made by the same rater. Indeed, the predictive agreement values for the OCI-CV are consistent with the performance

of other child-reported rating scales for related conditions (e.g., anxiety disorders, tic disorders; Caporino et al., 2017; Ricketts et al., 2018).

In comparisons to established benchmarks on the CY-BOCS (Skarphedinsson, et al., 2017), a 20% reduction in total score was optimal for determining treatment response, with a 65% reduction optimal for remission. Additionally, an OCI-CV total score ≤8 was optimal for determining clinical remission. The sensitivity (.43 - .81), specificity (.62 - .86), PPV (.57 - .65), NPV (.70 - .84), efficiency (.69 - .72) and Youden's J (.29 - .46) were slightly more variable across optimal benchmarks (relative to benchmarks with CGI scales), and remained lower than the sensitivity, specificity, PPV, NPV, and efficiency identified in benchmark studies using the CY-BOCS. In addition to rater training and/or composite parent and child report for the CY-BOCS, these differences may also be attributed to the weekly versus monthly assessment timeframe. For instance, a youth who experienced improvement from CBT in the past two weeks may see a noticeable reduction in the weekly CY-BOCS total score, but less improvement on the OCI-CV that inquires about symptoms over the past month.

Overall, fair to moderate agreement in treatment response and clinical remission classifications was observed between the OCI-CV, CY-BOCS, and CGI scales. The CGI-Improvement serves a global measure of improvement, and may be more attuned to the improvement of OCD symptoms that the youth and/or family find most distressing (e.g., reduced sexual obsessions). In comparison, the CY-BOCS is a comprehensive measure of OCD symptom severity, which may explain the higher agreement between the CY-BOCS and OCI-CV classifications compared the OCI-CV and CGI scales. Given the moderate classification agreement between benchmarks on the CY-BOCS and OCI-CV for treatment response (75% agreement) and clinical remission (73% agreement), the OCI-CV may serve as a useful stand-in for the CY-BOCS in applied practice—when clinician-rated scales cannot be administered.

When exploring factors associated with disagreement in classification, a lower baseline OCI-CV total score was associated with classification disagreement in determining treatment response with the CGI-Improvement and clinical remission with the CY-BOCS. Disagreement may be attributed to measurement differences, under-reporting of OCD symptoms by youth, limited treatment sensitivity of the OCI-CV compared to the CGI-Improvement, and/or limited specificity to detect remission compared to the CY-BOCS. Irrespective of these differences, when administering the OCI-CV at pre-treatment, it may be useful to orient the youth to the questionnaire to ensure full comprehension of items and symptoms. Indeed, there was one youth who reported having no OCD symptoms on the OCI-CV at pre-treatment, but had a clinician-rated CY-BOCS total score that met inclusion criteria. Thus, this initial orientation may reduce factors associated with classification disagreement and improve the precision of benchmark accuracy.

Despite study strengths (e.g., multi-site design, rigorous methodology, ongoing IE oversight to minimize rater drift), a few limitations remain. First, although this evaluation included participants on psychiatric medication, the focus of treatment was CBT. While most families prefer CBT to pharmacotherapy (Lewin, McGuire, Murphy, & Storch, 2014), it should be

noted that OCI-CV benchmarks may be different for monotherapy medication management. Second, this evaluation consisted of youth who participated in a clinical trial of CBT augmented with either d-cycloserine or placebo. The controlled nature of the study and participants' willingness to take an antibiotic medication may raise questions concerning the generalizability to a broader population of clinically-referred youth. Third, we did not collect parent-reported OCI-CV scores. As some reports suggest there to be high levels of parent-child disagreement in pediatric OCD (Canavera, Wilkins, Pincus, & Ehrenreich-May, 2009; Lewin, Peris, De Nadai, McCracken, & Piacentini, 2012), future research should examine whether optimal benchmarks would be different using parent-report.

Collectively, these findings provide optimal benchmarks for determining treatment response and clinical remission in pediatric OCD using a self-report scale, the OCI-CV. Although the CY-BOCS remains the gold standard assessment of OCD symptom severity, it is challenging to implement in a clinical practice due to limited resources (e.g., time, training). In the absence of administering a CY-BOCS, the OCI-CV can serve as an efficient alternative to monitor therapeutic progress that can be easily implemented. When using the OCI-CV, the optimal benchmarks for treatment response (≥20-25% reduction in total score) and clinical remission (a ≥55-65% reduction in total score or ≤6-8 total score cutoff) are offered to monitor therapeutic progress in evidence-based clinical care.

When implementing these findings in clinical practice, several important considerations are warranted. First, the OCI-CV should be used in the context of an evidence-based assessment (see Rapp et al., 2016). This typically includes a full diagnostic evaluation and clinician-administered OCD symptom severity scale. Thus, the OCI-CV should not be used as the sole means to assess an OCD diagnosis or categorize OCD symptom severity. Second, clinicians may rely on the CY-BOCS for pre-treatment, mid-treatment, and post-treatment administration; and utilize the OCI-CV as a more frequent measure of treatment response and clinical remission. This can be helpful as one approach to determine when might be the most opportune time to administer clinician assessments in treatment, thereby allocating scare therapeutic resources effectively. Finally, the OCI-CV should only be used as a substitute for the CY-BOCS when clinician-rated scales cannot be administered due to limited therapeutic resources (e.g., time, training).

Acknowledgements:

We acknowledge the contributions of Adam Lewin, Ph.D., Jane Mutch, Ph.D., Allison Kennel, ARNP, Nicole McBride, BS; Noah Berman, Ph.D., Alyssa Faro Kesley Ramsay, Ashley Brown, Andrew Mittelman, Abigail Stark, Allison Cooperman, and Angelina Gomez. This work was supported in part by grants 1R01MH093381 (Dr. Storch) and 5R01MH093402 (Dr. Geller) from the National Institute of Mental Health. The funding organization was not involved in the design or conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, and approval of the manuscript; or decision to submit the manuscript for publication.

References

Canavera KE, Wilkins KC, Pincus DB, & Ehrenreich-May JT (2009). Parent–child agreement in the assessment of obsessive-compulsive disorder. Journal of Clinical Child & Adolescent Psychology, 38, 909–915. 10.1080/15374410903258975 [PubMed: 20183673]

Caporino NE, Sakolsky D, Brodman DM, McGuire JF, Piacentini J, Peris TS, . . . Birmaher B (2017). Establishing Clinical Cutoffs for Response and Remission on the Screen for Child Anxiety Related

Emotional Disorders (SCARED). Journal of the American Academy of Child and Adolescent Psychiatry, 56(8), 696–702. 10.1016/j.jaac.2017.05.018 [PubMed: 28735699]

Enders CK (2010). Applied Missing Data Analysis. New York: The Guilford Press.

Foa EB, Coles M, Huppert JD, Pasupuleti RV, Franklin ME, & March J (2010). Development and validation of a child version of the obsessive compulsive inventory. Behavior Therapy, 41(1), 121–132. 10.1016/j.beth.2009.02.001 [PubMed: 20171333]

Geller DA, March J, & the AACAP Committee on Quality Issues. (2012). Practice Parameter for the Assessment and Treatment of Children and Adolescents With Obsessive-Compulsive Disorder. Journal of the American Academy of Child and Adolescent Psychiatry, 51(1), 98–113. 10.1016/j.jaac.2011.09.019 [PubMed: 22176943]

Guy W (1976). Clinical global impression scale. Rockville, MD: National Institute of Mental Health.

Johnco CJ, Salloum A, Lewin AB, & Storch EA (2015). Refining clinical judgment of treatment response and symptom remission identification in childhood anxiety using a signal detection analysis on the Pediatric Anxiety Rating Scale. Journal of Child and Adolescent Psychopharmacology, 25(9), 674–683. 10.1089/cap.2015.0102 [PubMed: 26579629]

Jones AM, De Nadai AS, Arnold EB, McGuire JF, Lewin AB, Murphy TK, & Storch EA (2013). Psychometric properties of the obsessive compulsive inventory: child version in children and adolescents with obsessive–compulsive disorder. Child Psychiatry & Human Development, 44(1), 137–151. 10.1007/s10578-012-0315-0 [PubMed: 22711294]

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, . . . Ryan N (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. Journal of the American Academy of Child and Adolescent Psychiatry, 36(7), 980–988. 10.1097/00004583-199707000-00021 [PubMed: 9204677]

Kraemer HC (1992). Evaluating medical tests: Objective and quantitative guidelines: Sage publications Newbury Park, CA:.

Lewin AB, McGuire JF, Murphy TK, & Storch EA (2014). Editorial Perspective: The importance of considering parent's preferences when planning treatment for their children–the case of childhood obsessive-compulsive disorder. Journal of Child Psychology and Psychiatry, 55, 1314–1316. 10.1111/jcpp.12344 [PubMed: 25346230]

Lewin AB, Peris TS, De Nadai AS, McCracken JT, & Piacentini J (2012). Agreement between therapists, parents, patients, and independent evaluators on clinical improvement in pediatric obsessive-compulsive disorder. Journal of Consulting and Clinical Psychology, 80, 1103–1107. 10.1037/a0029991 [PubMed: 22963592]

McGuire JF, Lewin AB, Geller DA, Brown A, Ramsey K, Mutch J, … Wilhelm S (2012). Advances in the treatment of pediatric obsessive–compulsive disorder: rationale and design for the evaluation of D-cycloserine with exposure and response prevention. Neuropsychiatry, 2(4), 291–300. 10.2217/npy.12.38

McGuire JF, Piacentini J, Lewin AB, Brennan EA, Murphy TK, & Storch EA (2015). A Meta-Analysis of Cognitive Behavior Therapy and Medication for Child Obsessive Compulsive Disorder: Moderators of Treatment Efficacy, Response, and Remission. Depression and Anxiety, 32(8), 580–593. 10.1002/da.22389 [PubMed: 26130211]

Ng MY, & Weisz JR (2016). Annual Research Review: Building a science of personalized intervention for youth mental health. Journal of Child Psychology and Psychiatry, and Allied Disciplines, 57(3), 216–236. 10.1111/jcpp.12470

Peugh JL, & Enders CK (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. Review of Educational Research, 74(4), 525–556.

Piacentini J, Bergman RL, Keller M, & McCracken J (2003). Functional impairment in children and adolescents with obsessive-compulsive disorder. Journal of Child and Adolescent Psychopharmacology, 13(2,Suppl), S61–S69. 10.1089/104454603322126359 [PubMed: 12880501]

Rapp A, Bergman R, Piacentini J, & McGuire JF (2016). Evidence-Based Assessment of Obsessive-Compulsive Disorder. Journal of Central Nervous System Disease, 8, 13–29. 10.4137/JCNSD.S38359 [PubMed: 27594793]

Ricketts EJ, McGuire JF, Chang S, Bose D, Rasch MM, Woods DW, … Wilhelm S (2018). Benchmarking Treatment Response in Tourette's Disorder: A Psychometric Evaluation and Signal Detection Analysis of the Parent Tic Questionnaire. Behavior Therapy, 49(1), 46–56. 10.1016/j.beth.2017.05.006 [PubMed: 29405921]

Rodriguez-Jimenez T, Piqueras JA, Lázaro L, Moreno E, Ortiz AG, & Godoy A (2016). Metric invariance, reliability, and validity of the Child Version of the Obsessive Compulsive Inventory (OCI-CV) in community and clinical samples. Journal of Obsessive-Compulsive and Related Disorders, 9, 1–8. 10.1016/j.jocrd.2016.01.003

Ruscio A, Stein D, Chiu W, & Kessler R (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. Molecular psychiatry, 15(1), 53–63. 10.1038/mp.2008.94 [PubMed: 18725912]

Scahill L, Riddle MA, McSwiggin-Hardin M, Ort SI, King RA, Goodman WK, … Leckman JF (1997). Children's Yale-Brown obsessive compulsive scale: reliability and validity. Journal of the American Academy of Child and Adolescent Psychiatry, 36(6), 844–852. 10.1097/00004583-199706000-00023 [PubMed: 9183141]

Skarphedinsson G, De Nadai AS, Storch EA, Lewin AB, & Ivarsson T (2017). Defining cognitive-behavior therapy response and remission in pediatric OCD: a signal detection analysis of the Children's Yale-Brown Obsessive Compulsive Scale. European Child and Adolescent Psychiatry, 26(1), 47–55. 10.1007/s00787-016-0863-0 [PubMed: 27209422]

Storch EA, Lewin AB, De Nadai AS, & Murphy TK (2010). Defining treatment response and remission in obsessive-compulsive disorder: a signal detection analysis of the Children's Yale-Brown Obsessive Compulsive Scale. Journal of the American Academy of Child and Adolescent Psychiatry, 49(7), 708–717. 10.1016/j.jaac.2010.04.005 [PubMed: 20610140]

Storch EA, Small BJ, McGuire JF, Murphy TK, Wilhelm S, & Geller DA (2017). Quality of Life in Children and Adolescents with Obsessive Compulsive Disorder. Journal of Child and Adolescent Psychopharmacology, 28(2), 104–110. 10.1089/cap.2017.0091 [PubMed: 28910139]

Storch EA, Wilhelm S, Sprich S, Henin A, Micco J, Small BJ, … Murphy TK (2016). Efficacy of augmentation of cognitive behavior therapy with weight-adjusted d-Cycloserine vs Placebo in pediatric obsessive-compulsive disorder: a randomized clinical trial. JAMA Psychiatry, 73(8), 779–788. 10.1001/jamapsychiatry.2016.1128 [PubMed: 27367832]

Highlights

- Evidence-based assessment is an important feature of evidence-based treatments
- Self-report scales are efficient tools to identify and quantify OCD symptom severity
- The OCI-CV is a psychometrically validated and commonly used measure in youth
- Benchmarks for treatment response and remission were identified on the OCI-CV
- The OCI-CV can be integrated into therapy to monitor clinical outcomes

Table 1.

Receiver Operator Curve and Quality Receiver Operator Curve Values for the Obsessive Compulsive Inventory-Child Version (OCI-CV) Percent Reduction in Total Scores Determining Treatment Response (N = 91)

Clinical Global Impression of Improvement							
Percent Reduction	**Sensitivity[a]**	**Specificity[b]**	**PPV[c]**	**NPV[d]**	**Efficiency[e]**	**Youden's J[f]**	**k(.05)[g]**
≥ 15	0.75	0.50	0.49	0.76	0.60	0.25	0.16
≥ 20	0.73	0.64	0.56	0.79	0.68	0.37	0.24
≥ 25	**0.69**	**0.71**	**0.60**	**0.78**	**0.70**	**0.40**	**0.27**
≥ 30	0.64	0.71	0.58	0.76	0.68	0.35	0.24
≤ 35	0.56	0.79	0.62	0.74	0.70	0.34	0.24
≤ 40	0.52	0.79	0.60	0.72	0.68	0.31	0.22
Children's Yale-Brown Obsessive Compulsive Scale (CYBOCS) total score reduction ≥35%							
Percent Reduction	**Sensitivity[a]**	**Specificity[b]**	**PPV[c]**	**NPV[d]**	**Efficiency[e]**	**Youden's J[f]**	**k(.05)[g]**
≥ 15	0.81	0.48	0.49	0.80	0.61	0.29	0.24
≥ 20	**0.81**	**0.62**	**0.57**	**0.84**	**0.69**	**0.43**	**0.36**
≥ 25	0.76	0.66	0.58	0.81	0.69	0.41	0.36
≥ 30	0.71	0.69	0.59	0.79	0.70	0.40	0.35
≥ 35	0.63	0.76	0.62	0.76	0.71	0.39	0.35
≥ 40	0.60	0.79	0.64	0.76	0.72	0.39	0.26

Note: Statistics for the cutoff with the highest quality of efficiency are in boldface type. PPV = Positive Predictive Value, NPV = Negative Predictive Value.

[a] Probability of exceeding the OCI-CV cutoff in youth meeting the treatment response criterion.

[b] Probability of not exceeding the OCI-CV cutoff in youth not meeting the treatment response criterion.

[c] Probability of meeting the treatment response criterion for those youth who exceeded the OCI-CV cutoff.

[d] Probability of not meeting the treatment response criterion for those youth who do not exceeded the OCI-CV cutoff.

[e] Probability that the OCI-CV cutoff and the dichotomous rating of treatment response agree.

[f] Youden's J measures optimal tradeoff of sensitivity and specificity at any single cutoff score.

[g] Weighted *K* statistic measuring quality of efficiency

Table 2.

Receiver Operator Curve and Quality Receiver Operator Curve Values for the Obsessive Compulsive Inventory-Child Version (OCI-CV) Percent Reduction in Total Scores Determining Clinical Remission (N = 91)

Clinical Global Impression of Severity							
Percent Reduction	**Sensitivity[a]**	**Specificity[b]**	**PPV[c]**	**NPV[d]**	**Efficiency[e]**	**Youden's J[f]**	**k(.05)[g]**
≥ 45	0.68	0.71	0.60	0.78	0.70	0.39	0.33
≥ 50	0.68	0.71	0.60	0.78	0.70	0.39	0.33
≥ 55	**0.63**	**0.76**	**0.63**	**0.77**	**0.71**	**0.40**	**0.34**
≥ 60	0.58	0.78	0.62	0.75	0.70	0.36	0.31
≥ 65	0.53	0.82	0.65	0.73	0.71	0.35	0.31
≥ 70	0.47	0.86	0.68	0.72	0.71	0.34	0.31
Children's Yale-Brown Obsessive Compulsive Scale (CYBOCS) total score reduction ≥55%							
Percent Reduction	**Sensitivity[a]**	**Specificity[b]**	**PPV[c]**	**NPV[d]**	**Efficiency[e]**	**Youden's J[f]**	**k(.05)[g]**
≥ 55	0.49	0.79	0.59	0.71	0.67	0.27	0.28
≥ 60	0.46	0.80	0.59	0.70	0.67	0.26	0.28
≥ 65	**0.43**	**0.86**	**0.65**	**0.70**	**0.69**	**0.29**	**0.31**
≥ 70	0.37	0.89	0.69	0.69	0.69	0.26	0.29
≥ 75	0.37	0.89	0.69	0.69	0.69	0.26	0.29
≥ 80	0.31	0.89	0.65	0.67	0.67	0.21	0.23

Note: Statistics for the cutoff with the highest quality of efficiency are in boldface type. PPV = Positive Predictive Value, NPV = Negative Predictive Value.

[a] Probability of exceeding the OCI-CV cutoff in youth meeting the clinical remission criterion.

[b] Probability of not exceeding the OCI-CV cutoff in youth not meeting the clinical remission criterion.

[c] Probability of meeting the clinical remission criterion for those youth who exceeded the OCI-CV cutoff.

[d] Probability of not meeting the clinical remission criterion for those youth who do not exceeded the OCI-CV cutoff.

[e] Probability that the OCI-CV cutoff and the dichotomous rating agree.

[f] Youden's J measures optimal tradeoff of sensitivity and specificity at any single cutoff score.

[g] Weighted *K* statistic measuring quality of efficiency

Table 3.

Receiver Operator Curve and Quality Receiver Operator Curve Values for the Obsessive Compulsive Inventory-Child Version (OCI-CV) Absolute Score Cutoffs for Determining Clinical Remission (N = 91)

Clinical Global Impression of Severity							
Total Score Cutoffs	**Sensitivity[a]**	**Specificity[b]**	**PPV[c]**	**NPV[d]**	**Efficiency[e]**	**Youden's J[f]**	**k(.05)[g]**
≤ 4	0.42	0.78	0.54	0.68	0.64	0.20	0.18
≤ 6	**0.63**	**0.74**	**0.60**	**0.76**	**0.70**	**0.37**	**0.32**
≤ 8	0.79	0.60	0.55	0.82	0.67	0.39	0.31
≤ 10	0.84	0.49	0.51	0.83	0.62	0.33	0.25
Children's Yale-Brown Obsessive Compulsive Scale total score cutoff ≤11							
Total Score Cutoffs	**Sensitivity[a]**	**Specificity[b]**	**PPV[c]**	**NPV[d]**	**Efficiency[e]**	**Youden's J[f]**	**k(.05)[g]**
≤ 4	0.41	0.85	0.63	0.70	0.68	0.26	0.28
≤ 6	0.56	0.83	0.67	0.75	0.73	0.39	0.41
≤ 8	**0.74**	**0.71**	**0.62**	**0.82**	**0.72**	**0.46**	**0.45**
≤ 10	0.82	0.60	0.56	0.84	0.68	0.42	0.39
≤ 12	0.90	0.58	0.57	0.90	0.70	0.47	0.44
≤ 14	0.90	0.42	0.49	0.87	0.61	0.32	0.28

Note: Statistics for the cutoff with the highest quality of efficiency are in boldface type. PPV = Positive Predictive Value, NPV = Negative Predictive Value.

[a] Probability of exceeding the OCI-CV cutoff in youth meeting the clinical remission criterion.

[b] Probability of not exceeding the OCI-CV cutoff in youth not meeting the clinical remission criterion.

[c] Probability of meeting the clinical remission criterion for those youth who exceeded the OCI-CV cutoff.

[d] Probability of not meeting the clinical remission criterion for those youth who do not exceeded the OCI-CV cutoff.

[e] Probability that the OCI-CV cutoff and the dichotomous rating agree.

[f] Youden's J measures optimal tradeoff of sensitivity and specificity at any single cutoff score.

[g] Weighted *K* statistic measuring quality of efficiency

Table 4.

Agreement and Disagreement in Treatment Response and Clinical Remission Classification Methods (N=91)

	Agreement (N, %)	Disagreement (N, %)	Kappa	Agreement[d]	χ^2
Agreement and Disagreement Between Treatment Response Classifications[a]					
OCI-CV & CGI-Improvement	63 (69%)	28 (31%)	.25	Fair	8.20**
OCI-CV and CY-BOCS	68 (75%)	23 (25%)	.42	Moderate	16.31***
CY-BOCS and CGI-I	76 (84%)	15 (16%)	.56	Moderate	35.37***
Agreement and Disagreement Between Clinical Remission Classifications Using Percent Reduction[b]					
OCI-CV and CGI-Severity	67 (74%)	24 (26%)	.33	Fair	10.83***
OCI-CV and CY-BOCS	63 (69%)	28 (31%)	.31	Fair	9.31**
CY-BOCS and CGI-S	73 (80%)	18 (20%)	.54	Moderate	32.13***
Agreement and Disagreement Between Clinical Remission Classifications Using Cutoff Scores[c]					
OCI-CV and CGI-S	65 (72%)	26 (28%)	.30	Fair	9.05**
OCI-CV and CY-BOCS	66 (73%)	25 (27%)	.45	Moderate	18.49***
CY-BOCS and CGI-S	69 (76%)	22 (24%)	.47	Moderate	26.39***

Note: OCI-CV = Obsessive Compulsive Inventory- Child Version, CGI-Improvement = Clinical Global Impression of Improvement, CY-BOCS = Children's Yale Brown Obsessive Compulsive Scale, CGI-Severity = Clinical Global Impression of Severity.

[a]Classifications included a ≥25% reduction in OCI-CV total score, and a ≥35% reduction in CY-BOCS total score

[b]Classifications included a ≥55% reduction in OCI-CV total score, and a ≥55% reduction in CY-BOCS total score.

[c]Classification included a total score cutoff of ≤6 on the OCI-CV and a total score cutoff of ≤11 on the CY-BOCS

[d]Landis and Koch (1977) consider < 0 as indicating no agreement and 0–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1 as almost perfect agreement.

*p<.05,

** $p<.01$, *** $p<.001$.