Published in final edited form as:

J Am Acad Child Adolesc Psychiatry. 2013 January ; 52(1): 57–67. doi:10.1016/j.jaac.2012.10.006.

Defining Treatment Response and Remission in Child Anxiety: Signal Detection Analysis Using the Pediatric Anxiety Rating Scale

Dr. Nicole E. Caporino, Ph.D.,
Temple University

Mr. Douglas M. Brodman, M.A.,
Temple University

Dr. Philip C. Kendall, Ph.D.,
Temple University

Dr. Anne Marie Albano, Ph.D.,
Columbia University Medical Center

Dr. Joel Sherrill, Ph.D.,
Division of Services and Intervention Research at the National Institute of Mental Health (NIMH)

© 2013 American Academy of Child and Adolescent Psychiatry

Correspondence to Nicole E. Caporino, Ph.D., Department of Psychology, Temple University, Weiss Hall, 1701 N. 13th Street, Philadelphia, PA 19122; nicole.caporino@temple.edu.

Sertraline and matching placebo were supplied free of charge by Pfizer.

Views expressed within this article represent those of the authors and are not intended to represent the position of NIMH, the National Institutes of Health (NIH), or the Department of Health and Human Services.

Disclosure: Dr. Kendall has received research support from NIMH. He has received honoraria from professional societies for speaking at conventions. He has received royalties from Guilford Press, Ericsson, Workbook Publishing, and Oxford University Press. Dr. Albano has received research support from NIMH. She has received honoraria from the American Psychological Association. She has served as a consultant to Brackett Global. She has received royalties from Oxford University Press. Dr. Piacentini has received grant or research support from NIMH, the Tourette Syndrome Association, and Otsuka Pharmaceuticals. He has reviewed survey materials related to the 2010 World Contraception Day media report for Bayer Schering Pharma. He has received book royalties from Guilford Press and Oxford University Press. He is a co-author of assessment tools, none of which are commercially published and therefore no royalties have been received. He has received speaking honoraria/travel support from the Tourette Syndrome Association and the International Obsessive Compulsive Disorder Foundation. Dr. Sakolsky has received research support from NIMH and the National Alliance for Research on Schizophrenia and Depression (NARSAD). She has received honoraria from the American Academy of Child and Adolescent Psychiatry for participation in the 37th Annual Review Course in Child and Adolescent Psychiatry and Training for the Oral Exams. Dr. Birmaher has received research support from NIMH. He has received royalties from Random House. Dr. Compton has received research support from NIMH. He has served as a consultant to Shire Pharmaceuticals. He has provided expert forensic testimony for mental health needs of children in high-conflict families. Dr. Rynn has received research support from NIMH, the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), Eli Lilly and Co., Pfizer, Merck, and Shire. She has served as a consultant to Shire. She has received royalties from American Psychiatric Publishing. Dr. McCracken has received research support from Seaside Therapeutics, Roche, and Otsuka. He has served as a consultant to BioMarin, PharmaNet, Roche, and Novartis. He has received research study drug from Shire. Dr. Gosch has received royalties from Springer Publishing Company. Dr. Keeton has received research support from NIMH and the Johns Hopkins Urban Health Institute. Dr. March has received research support from NIMH, the National Institute on Drug Abuse (NIDA), NARSAD, and Pfizer. He has served as a consultant to Atentiv, Bristol-Myers Squibb, Eli Lilly and Co., Pfizer, and Widay Pharmaceuticals. He has served on scientific advisory boards of Eli Lilly and Co., Pfizer, and Shire. He has served on the Data Safety Monitoring Boards (DSMB) of Eli Lilly and Co., NIDA, and Pfizer. He has received royalties from Guilford Press, MultiHealth Systems, and Oxford University Press. He retains equity in MedAvante. He has provided expert forensic consultation to DLA Piper. Dr. Walkup has received grant or research support from the Tourette Syndrome Association. He has served as a consultant to Shire. He has received free medication and placebo from Eli Lilly and Co., Pfizer, and Abbott for NIH-funded studies. He has served on the advisory board and speakers' bureau of the Tourette Syndrome Association. He has received royalties from Guilford Press and Oxford University Press. He has received honorarium for an Educational Meeting from the Tourette Syndrome Association. He also has received travel support and honoraria for paid and unpaid activities from the Tourette Syndrome Association including an unpaid position on the Medical Advisory Board. Drs. Caporino, Sherrill, and Ginsburg, and Mr. Brodman report no biomedical financial interests or potential conflicts of interest.

Dr. John Piacentini, Ph.D.,
Semel Institute for Neuroscience and Human Behavior at the University of California Los Angeles

Dr. Dara Sakolsky, M.D., Ph.D.,
Western Psychiatric Institute and Clinic at the University of Pittsburgh Medical Center

Dr. Boris Birmaher, M.D.,
Western Psychiatric Institute and Clinic at the University of Pittsburgh Medical Center

Dr. Scott N. Compton, Ph.D.,
Duke University Medical Center

Dr. Golda Ginsburg, Ph.D.,
Johns Hopkins University School of Medicine

Dr. Moira Rynn, M.D.,
Columbia University Medical Center

Dr. James McCracken, M.D.,
Semel Institute for Neuroscience and Human Behavior at the University of California Los Angeles

Dr. Elizabeth Gosch, Ph.D.,
Philadelphia College of Osteopathic Medicine

Dr. Courtney Keeton, Ph.D.,
Johns Hopkins University School of Medicine

Dr. John March, M.D., M.P.H., and
Duke University Medical Center

Dr. John T. Walkup, M.D.
Weill Cornell Medical College

Abstract

Objective—To determine optimal Pediatric Anxiety Rating Scale (PARS) percent reduction and raw score cut-offs for predicting treatment response and remission among children and adolescents with anxiety disorders.

Method—Data were from a subset of youth ($N = 438$; 7–17 years of age) who participated in the Child/Adolescent Anxiety Multimodal Study (CAMS), a multi-site, randomized controlled trial that examined the relative efficacy of cognitive-behavioral therapy (CBT; *Coping Cat*), medication (sertraline [SRT]), their combination, and pill placebo for the treatment of separation anxiety disorder, generalized anxiety disorder, and social phobia. The clinician-rated PARS was administered pre- and posttreatment (delivered over 12 weeks). Quality receiver operating characteristic methods assessed the performance of various PARS percent reductions and absolute cut-off scores in predicting treatment response and remission, as determined by posttreatment ratings on the Clinical Global Impression scales and the Anxiety Disorders Interview Schedule for *DSM-IV*. Corresponding change in impairment was evaluated using the Child Anxiety Impact Scale.

Results—Reductions of 35% and 50% on the six-item PARS optimally predicted treatment response and remission, respectively. Post-treatment PARS raw scores of 8 to 10 optimally predicted remission. Anxiety improved as a function of PARS-defined treatment response and remission.

Conclusions—Results serve as guidelines for operationalizing treatment response and remission in future research and in making cross-study comparisons. These guidelines can facilitate translation of research findings into clinical practice.

Keywords

anxiety; treatment; measurement

Research has supported the efficacy of selective-serotonin reuptake inhibitors (SSRIs) and cognitive-behavioral therapy (CBT) for the treatment of anxiety disorders in children and adolescents.^{1–5} Researchers have typically evaluated outcome by examining change in severity ratings on the Anxiety Disorders Interview Schedule for *DSM-IV*—Child and Parent Versions (ADIS-IV-C/P)⁶ and/or the Clinical Global Impressions Scales (CGI).⁷ ADIS-IV-C/P clinical severity ratings are specific to individual diagnoses and are assigned after interviews conducted with the child and the child's parent(s). CGI ratings represent global assessments of severity (CGI-Severity) and improvement (CGI-Improvement), and are assigned based on all information available from the assessment. CGI ratings typically refer to overall psychopathology; however, some studies have used them to evaluate target disorders.⁵ Although these measures can be seen as the “gold standard,” they can be time-consuming to administer and require lengthy training. Thus, they are less central to “patient-focused research” aimed at assessing treatment response for individual cases in clinical practice.^{8,9}

The Pediatric Anxiety Rating Scale (PARS)¹⁰ has been used as a dimensional measure of treatment efficacy.⁵ The PARS is a clinician-rated measure of symptom severity and associated impairment that targets generalized anxiety disorder (GAD), social phobia (SoP), and separation anxiety disorder (SAD). The PARS has acceptable psychometric properties and is sensitive to change in CBT and pharmacological treatment.^{5,10,11} The comprehensive nature of the PARS is appealing in light of symptom overlap and high rates of comorbidity across anxiety disorders.^{12,13} The PARS is time-efficient (taking approximately 20–30 minutes to complete). Similar to other interview-based rating scales that assess severity and response to treatment, such as the Children's Depression Rating Scale—Revised¹⁴ and the Children's Yale–Brown Obsessive Compulsive Scale (CY-BOCS)¹⁵, the PARS may be feasible for use in routine clinical care.

Despite the potential merits of the PARS, there are currently few guidelines for its use in determining treatment response and remission. “Treatment response” has been defined as improvement of sufficient magnitude such that the individual is no longer fully symptomatic but may continue to evidence more than minimal symptoms.¹⁶ Treatment response is often operationalized as a significant reduction in a measure of symptom severity and/or functional impairment. “Remission” has been defined as the absence or near absence of symptoms after treatment¹⁷ and may be important in the treatment of childhood disorders given the impact of residual symptoms on development.¹⁸ Relative to treatment response, remission is a more conservative standard and has been operationalized using binary measures of diagnostic status or dichotomized ratings on dimensional measures of global functioning, which correspond to youth being “disorder free.”¹⁹ Both treatment response and remission should be defined *a priori* and measured using multiple sources of information.¹⁶ Multiple definitions of remission were used in the Child/Adolescent Anxiety Multimodal Study (CAMS)⁵, including CGI-Severity ratings corresponding to minimal or no symptoms and the loss of all targeted diagnoses on the ADIS-IV-C/P.¹⁹ Although the latter definition had been used in prior studies of CBT for child anxiety,²⁰ it allows residual symptoms at the subclinical level rather than requiring the absence or near absence of symptoms. Despite the need for a consensus definition, establishing criteria for using the PARS to determine treatment response and remission could facilitate comparisons across clinical trials and improve treatment guidelines.

We used CAMS data to delineate criteria for using PARS scores to identify clinically meaningful change, as determined by reliable evaluators using acceptable measures. Specifically, we sought to identify the optimal PARS percent reduction and absolute cut-off scores for predicting treatment response (CGI-Improvement rating of 1 or 2) and clinical remission (CGI-Severity rating of 1 or 2; loss of all targeted diagnoses on the ADIS-IV-C/P).

METHOD

Participants

Participants were CAMS youth⁵ who were administered all measures of interest before and after treatment. This subset of CAMS participants comprised 438 youth (51% female and 49% male) with a principal diagnosis of SAD (31%), GAD (48%), and/or SoP (49%). They ranged in age from 7 to 17 years (mean = 10.72, SD = 2.80) and were recruited with a parent across six university-based outpatient clinics. Youth were randomly assigned to one of four treatment conditions: sertraline (SRT; n = 115), cognitive-behavioral therapy (CBT; n = 132), a combination of these two (COMB; n = 127), or pill placebo (n = 64), representing 86% of SRT, 95% of CBT, 91% of COMB, and 84% of pill placebo conditions in CAMS. The racial distribution of the sample was 81% white (n = 353), 8% African American (n = 35), 3% Asian (n = 11), 1% Native Hawaiian/Other Pacific Islander (n = 5), and 2% biracial (n = 8). With regard to ethnicity, 11% (n = 48) of the sample identified as Hispanic/Latino. Demographics of the current sample (N = 438) did not differ significantly from the sample in CAMS (N = 488). CAMS inclusion and exclusion criteria have been described²¹ and the sample evidenced high comorbidity; attention-deficit/hyper-activity disorder (10.04%; n = 49), oppositional defiant disorder (9.43%; n = 46), and obsessive-compulsive disorder (8.61%; n = 42) were most common outside of targeted anxiety disorders.¹²

Procedures

Procedures were approved by the Institutional Review Board at each of the six sites where data were collected. Measures were administered by independent evaluators (IEs) blind to treatment condition. In addition to having prior experience with administration of the primary measures, IEs were trained to a pre-specified reliability criterion and monitored for drift. IEs participated in weekly on-site supervision and biweekly cross-site calibration conference calls during the study.

Measures

Pediatric Anxiety Rating Scale (PARS).¹⁰—The PARS is a clinician-rated measure of anxiety severity in children and adolescents. The PARS consists of a checklist of 50 anxiety symptoms (encompassing SAD, SoP, and GAD) and seven global items that are administered to the child and parent(s) together. Global items are each rated on a six-point (0–5) scale and reflect the number of symptoms present, their frequency, the severity of anxiety feelings, the severity of physical symptoms of anxiety, overall avoidance of anxiety-provoking situations, and anxiety-related interference with functioning at and outside of the home. Acceptable reliability and validity data have been demonstrated.¹⁰ In CAMS, six global items were summed to produce a total score with a possible range of 0 to 30; the item assessing the number of symptoms present (per checklist) was not included in the total score due to concern about overlap with the item assessing the frequency of symptoms. PARS total scores ranged from 7 to 29 at baseline. Inter-rater reliability in CAMS was determined based on a review of 10% of videotaped assessments conducted before and after treatment (Pearson's $r = 0.85$).⁵

Clinical Global Impressions (CGI).⁷—The CGI Scales are commonly used in clinical trials and provide clinician ratings of the severity of psychopathology (CGI-Severity) and overall improvement (CGI-Improvement). The CGI-Severity scale ranges from 1 (no illness) to 7 (extremely severe), whereas the CGI-Improvement scale ranges from 1 (very much improved) to 7 (very much worse). Consistent with Walkup *et al.*,⁵ a CGI-Improvement score of 1 (very much improved) or 2 (much improved) was used to designate treatment response. Consistent with Ginsburg *et al.*,¹⁹ a CGI-Severity score of 1 (not at all ill) or 2 (borderline ill) was used to designate remission. These operational definitions have been used to evaluate outcomes²² and correspond closely to early conceptualizations of response and remission.¹⁷

Anxiety Disorders Interview Schedule for DSM-IV–Child and Parent Versions (ADIS-IV-C/P).⁶—The ADIS-IV-C/P is a clinician-administered semi-structured diagnostic interview that assesses major DSM-IV anxiety disorders and associated psychopathology (e.g., disruptive behavior disorders) in school-aged children and adolescents. Composite diagnoses are determined using information gathered from parent and child interviews, and are based on the presence of core symptoms and a clinical severity rating of 4 or greater (on a scale of 0–8). Consistent with prior research,²⁰ loss of targeted diagnoses on the ADIS-IV-C/P was used as a second operational definition of remission that is in line with the use of diagnostic status to determine need for treatment and inclusion in controlled trials. The ADIS-IV-C/P has excellent psychometric properties.^{23,24} Based on a review of 10% of videotaped pre- and posttreatment assessments, interrater reliability for diagnostic status (intraclass correlation coefficient) in CAMS ranged from 0.82 to 0.88.

Child Anxiety Impact Scale-Revised (CAIS-R/P).²⁵—The CAIS-R/P²⁵ (Langley, Peris, Piacentini, unpublished material, 2012) is a 27-item parent-report measure of anxiety-related interference in social activities, school, and home/family functioning. The CAIS-R/P has good internal consistency, construct validity, and discriminant validity. Possible scores range from 0 to 81; posttreatment scores in this sample ranged from 0 to 79.

Data Analysis Plan

Consistent with analyses in previous studies,^{26,16} receiver operating characteristic (ROC) methods^{27,28,29} were used to assess the performance of various PARS percent reduction and absolute cut-off scores at predicting response (i.e., CGI-Improvement of 1 or 2) and remission (i.e., CGI-Severity of 1 or 2; loss of all targeted diagnoses on ADIS-IV-C/P). PARS scores were divided into cut-offs based on percent reduction (in 5% intervals) and raw total scores. For each PARS (i.e., test) cutoff, analyses were conducted to examine the sensitivity (probability that youth meeting the gold standard criterion will exceed the test cut-off), specificity (probability that youth not meeting the gold standard criterion will not exceed the test cut-off), positive predictive value (probability that youth exceeding the test cut-off will meet the gold standard criterion), negative predictive value (probability that youth not exceeding the test cut-off will not meet the gold standard criterion), and efficiency (probability that the test and the gold standard will agree). Youden's J ($J = \text{sensitivity} + \text{specificity} - 1$)³⁰ is also reported as a measure of optimal tradeoff between sensitivity and specificity at a single cut-off point. Youden's J ranges from −1 (indicating no discrimination) to +1 (indicating perfect discrimination). ROC curves depict the balance between sensitivity (plotted on the abscissa) and specificity (plotted on the ordinate) in evaluating test performance.

Quality receiver operating characteristic methods (QROC)^{31,32} were used. QROC methods were developed in response to concerns about the application of traditional ROC analyses to biomedical and behavioral research given error in measurement of the criterion variable

(“gold standard”). QROC methods involve using specific forms of weighted κ statistics that assess the quality of various ROC statistics.^{31,32} This study used the $\kappa(0.0)$, $\kappa(0.5)$, and $\kappa(1.0)$ statistics, which measure the quality of specificity, quality of efficiency, and quality of sensitivity, respectively. For each of these statistics, a value of 0.00 indicates that the property cannot be differentiated from chance in assessing responder/remitter status using the PARS cut-off, whereas a value of 1.00 indicates perfect assessment of responder/remitter status. These values can be plotted on a QROC plane (akin to the traditional ROC curve), with $\kappa(1.0)$ on the ordinate and $\kappa(0.0)$ on the abscissa. The top right apex of the curve (as opposed to the top left of traditional ROC curves) indicates maximal sensitivity and specificity. For the current study, priority was given to maximizing efficiency or accuracy of classification. Thus, optimal cut-offs were selected using $\kappa(0.5)$ to minimize false-positive and false-negative results simultaneously. All analyses were conducted for the total sample followed by each condition and principal diagnosis, using the six-item PARS. Analyses were repeated with the five-item PARS used in some medication trials; results were comparable and are available from the first author.

RESULTS

Missing Data

Data were analyzed for participants with both pre-and posttreatment measures. For these analyses, there were no data missing for the PARS, CGI scales, or ADIS-IV-C/P. For the CAIS-R/P, fewer than 10% of items were omitted for each participant and scores were calculated by averaging responses across the items that make up each scale. Analyses were repeated using the CAMS intent-to-treat sample and multiple imputation procedures for addressing missing data due to attrition, which was minimal. Results were similar and thus, not included in this report.

Descriptive Statistics

Among those participants who completed a week 12 assessment, the average pretreatment PARS score across all treatment conditions was 19.22 (SD = 4.15) and the average posttreatment score was 9.49 (SD = 6.6). A dependent-samples *t* test for PARS total score reduction was significant ($t[437] = -31.40$, $p < .001$, Cohen's $d = 1.76$). The average percent reduction in PARS total scores was 51% (SD = 33%).

The average pretreatment CGI-Severity rating was 5.03 (SD = 0.73) and the average posttreatment rating was 2.94 (SD = 1.47). A dependent samples *t* test for overall CGI-Severity reduction was significant ($t[437] = 30.84$, $p < .001$, Cohen's $d = 1.81$). The average percent reduction in CGI-Severity scores was 58% (SD = 28%). At posttreatment, approximately 65% of the total sample met criteria for treatment response based on CGI-Improvement ratings, 46% met criteria for remission based on CGI-Severity ratings, and 53% met criteria for remission based on ADIS-IV-C/P diagnostic status.

Neither baseline PARS scores nor baseline CGI-Severity scores were correlated with PARS percent reductions (suggesting that the present findings apply across baseline severity levels). PARS percent reductions were significantly associated with dichotomized CGI-Improvement ratings ($r_{pb} = 0.74$, $p < .001$) and with dichotomized CGI-Severity ratings ($r_{pb} = 0.74$, $p < .001$). The correlation between PARS percent reductions and ADIS-IV-C/P diagnostic status was also significant ($r_{pb} = .72$, $p < .001$).

Predicting Treatment Response With PARS Percent Reduction

Table 1 presents the series of PARS percent reduction cut-offs used to predict dichotomized CGI-Improvement ratings, with ROC statistics for each cut-off. Maximum efficiency for

predicting response was found at a cut-off of 35%. At this cut-off, predictive value of a positive test and predictive value of a negative test are each .89, indicating false-positive and false-negative rates of approximately 11%. Analyses did not reveal meaningful differences by principal diagnosis; quality of efficiency was consistently highest at a 35% reduction.

When analyses were conducted separately for each condition, a 35% reduction cut-off was optimal for SRT (efficiency = 0.94, $\kappa[0.5] = 0.87$) and CBT (efficiency = 0.87, $\kappa[0.5] = 0.72$). A 30% reduction cut-off was optimal for combination treatment (efficiency = 0.94, $\kappa[0.5] = 0.70$). For the placebo condition, which consisted of fewer participants and showed lower response rates than the active treatment conditions, a 55% reduction optimally predicted treatment response (efficiency = 0.89, $\kappa[0.5] = 0.72$).

Predicting Remission With PARS Percent Reduction

PARS percent reduction cut-offs were used to predict dichotomized CGI-Severity ratings. Maximum efficiency (efficiency = .80, $\kappa[0.5] = 0.59$) for predicting clinical improvement was found at a cut-off of 50%, with cut-offs of 45% (efficiency = 0.79, $\kappa[0.5] = 0.58$) and 55% (efficiency = 0.79, $\kappa[0.5] = 0.59$) very close in terms of efficiency ratings. Among these similar cut-offs, the 50% reduction cut-off shows the optimal tradeoff between sensitivity (0.81) and specificity (0.78), with acceptable levels of positive and negative predictive power (0.82 and 0.77, respectively). Results were consistent when analyses were repeated with clinical remission defined as loss of all targeted diagnoses (i.e., GAD, SoP, and/or SAD) on the ADIS-IV-C/P; maximum efficiency was again found at a cut-off of 50% reduction, albeit with slightly higher sensitivity (0.87), specificity (0.83), positive predictive value (0.86), and negative predictive value (0.85). ROC statistics for PARS percent reduction cut-offs used to predict loss of ADIS-IV-C/P diagnoses are displayed in Table 2. Analyses conducted separately for each principal diagnosis yielded similar results, with 45% to 50% reduction cut-offs consistently identified as optimal.

Analyses were repeated for each condition. A 50% reduction cut-off optimally predicted loss of ADIS-IV-C/P diagnoses for all three active treatments: CBT (efficiency = 0.80, $\kappa[0.5] = 0.59$), SRT (efficiency = 0.89, $\kappa[0.5] = 0.74$), and combination (efficiency = 0.90, $\kappa[0.5] = 0.74$). For the placebo condition, a 55% reduction cut-off optimally predicted loss of ADIS-IV-C/P diagnoses (efficiency = 0.89, $\kappa[0.5] = 0.72$).

A 50% reduction cut-off optimally predicted remission, defined using the CGI-Severity, for SRT (efficiency = 0.90, $\kappa[0.5] = 0.79$) and combination treatment (efficiency = 0.80, $\kappa[0.5] = 0.51$). A 65% reduction cut-off was optimal for CBT (efficiency = 0.85, $\kappa[0.5] = 0.66$), although ROC statistics were also good at the 50% reduction cut-off (efficiency = 0.80, $\kappa[0.5] = 0.66$). A 55% reduction cut-off was optimal for the placebo condition (efficiency = 0.89, $\kappa[0.5] = 0.69$).

Predicting Remission With PARS Total Scores

When remission was determined using dichotomized CGI-Severity ratings, maximal efficiency (efficiency = 0.90, $\kappa[0.5] = 0.80$) was found for a PARS raw score cut-off of 8, with strong sensitivity (0.93), specificity (0.88), positive predictive value (0.98), and negative predictive value (0.94) at this cut-off. When analyses were repeated with clinical remission defined as loss of all targeted diagnoses (i.e., GAD, SoP, and/or SAD) on the ADIS-IV-C/P, maximal efficiency was found for a raw score cut-off of 10. ROC statistics for PARS raw score cut-offs used to predict loss of ADIS-IV-C/P diagnoses are displayed in Table 3. Results were similar when analyses were conducted separately for each principal diagnosis; cut-offs of 8 to 10 consistently yielded maximal efficiency.

When analyses were conducted separately for each condition, a raw score cut-off of 8 optimally predicted remission, defined using the CGI-Severity, in the SRT (efficiency = 0.94, $\kappa[0.5] = 0.88$), CBT (efficiency = 0.86, $\kappa[0.5] = 0.72$), and placebo (efficiency = 0.89, $\kappa[0.5] = 0.67$) conditions. Efficiency was equally strong at a cut-off of 7 in the placebo condition. For combination treatment, the optimal cut-off score was 9 (efficiency = 0.93, $\kappa[0.5] = 0.84$).

When remission was defined as loss of diagnoses on the ADIS-IV-C/P, a raw score cut-off of 8 was optimal for SRT (efficiency = 0.91, $\kappa[0.5] = 0.82$) and CBT (efficiency = 0.81, $\kappa[0.5] = 0.61$). Efficiency was equally strong at a cut-off of 9 in the SRT condition. A raw score cut-off of 9 was optimal for combination treatment (efficiency = 0.90, $\kappa[0.5] = 0.76$) and placebo (efficiency = 0.84, $\kappa[0.5] = 0.61$).

Comparison of Response and Remission for PARS

Figure 1 shows the $\kappa(0.5)$ metric over the series of PARS percent reduction cut-offs for predicting response and remission using the total sample. Response shows maximal $\kappa(0.5)$ at 35% reduction. Remission (as determined using the CGI-Severity and the ADIS-IV-C/P) shows maximal $\kappa(0.5)$ at 50% reduction.

Functional Outcomes

Change in anxiety-related impairment in functioning was assessed using the CAIS-R/P. For participants classified as treatment responders using the PARS 35% reduction cut-off, effect sizes were Cohen's $d = 1.09$, 1.22, and 1.16 for social school, and home/family domains, respectively. The effect size for the CAIS-R/P total score among treatment responders was Cohen's $d = 1.47$. For participants identified as remitters using the PARS 50% reduction cut-off, effect sizes were somewhat higher, with Cohen's $d = 1.24$, 1.32, and 1.23 for social, school, and home/family domains, respectively. The effect size for the CAIS-R/P total score among remitters was Cohen's $d = 1.54$. All of these differences were significant at the $p < .01$ level.

DISCUSSION

Signal detection analysis of the six-item PARS, a relatively brief, clinician-rated measure of anxiety symptoms across disorders, revealed that a 35% reduction in total scores from pre-to posttreatment best predicted treatment response. A 50% reduction in PARS best predicted remission. Posttreatment PARS absolute cut-offs of 8 and 10 showed the strongest association with the "gold standard" measures of remission used in CAMS. Furthermore, each of the percent reduction and absolute cut-offs identified as optimal was associated with significant reductions in parent-reported impairments in functioning, providing further support that the cut-offs are meaningful. When analyses were repeated for each condition separately, results were similar across active treatments (i.e., SRT, CBT, and COMB).

The present results were consistent with a signal detection analysis of the CY-BOCS, a measure of obsessive-compulsive symptom severity after which the PARS was modeled. A 25% reduction on CY-BOCS scores was optimal for predicting treatment response²⁶ versus a 35% reduction on the PARS. For both measures, a 50% reduction was optimal for predicting remission using CGI-Severity. In line with notions of treatment response and remission,¹⁷ the threshold for remission was substantially higher than the threshold for treatment response in both studies. Thus, a treatment response does not necessarily indicate high levels of functioning at posttreatment.

Although the optimal PARS percent reduction cut-off was the same regardless of which operational definition of remission was used (CGI-S of 1–2 or loss of targeted ADIS-IV-C/P

diagnoses), absolute cut-off scores were slightly different (8 for CGI-defined remission and 10 for ADIS-defined remission). It may be that the CGI-Severity criterion is more stringent than the ADIS-IV-C/P criterion, which does not necessarily correspond to a near symptom-free state. This discrepancy underscores the need for a consensus definition of remission for youth anxiety disorders.

The six-item PARS absolute cut-offs (8–10) that best predicted remission outcomes were lower than the five-item PARS cut-off of 11.5 previously found to discriminate youth with and without anxiety disorders.³³ Methodological variations may explain these differences. Whereas comparison groups in the current study were youth following CAMS treatments (i.e., remitters versus non-remitters), Ginsburg *et al.*³³ compared youth with anxiety disorders recruited for a medication trial¹¹ to youth without anxiety disorders.^{34,35} The sample of youth without anxiety was relatively small and perhaps unrepresentative (i.e., included children of parents with anxiety).

Among the strengths of this study are that assessments were conducted by reliable IEs, multiple operational definitions of remission were used, and the sample included youth whose treatment was found to be efficacious.⁵ However, potential limitations are noted. First, inclusion and exclusion criteria may limit generalizability to the broader population of clinic-referred youth. However, empirical studies comparing youth treated through university-based research clinics to youth treated through community-based service clinics have shown minimal differences on measures of internalizing problems,³⁶ and youth with the most common externalizing comorbidities were included in CAMS. Second, CGI-Improvement and CGI-Severity ratings were made using all available information and may have been influenced by the PARS. Also, inter-rater reliability was not calculated for the CGI scales (although CGI ratings were provided by the same IEs who were reliable on the ADIS-IV-C/P and the PARS). Finally, analyses were carried out using those cases in which all measures were administered before and after treatment; results are to be generalized only to youth with anxiety who complete treatment.

The findings have implications for research and practice. For many childhood disorders, there are inconsistencies in how treatment successes are identified.^{37–39} By establishing a PARS percent reduction and/or absolute score cut-off(s) that denotes treatment response and remission, some confusion can be averted. Providing PARS cut-offs may improve the “uptake” of research findings by assisting clinicians in assessing the progress of individual patients against the standard of outcomes reported in clinical trials. Specifically, cut-offs can inform the decision to augment treatment with additional services. For example, if a patient who receives medication or CBT does not experience a 35% symptom reduction in 12 weeks, his/her clinician may recommend consultation with an additional service provider. Although cut-offs may be best viewed as a single source of information, learning and using the PARS requires minimal effort and the benefit of using cut-offs to guide treatment planning is underscored by evidence that therapists are often not alert to treatment failure.^{40,41} Future research should establish empirically based guidelines for identifying extended remission without relapse.¹⁷ In addition, cut-offs for parent- and child-report measures might prove useful in settings in which multiple assessments are not possible.

Acknowledgments

This research was supported by NIMH grants U01 MH064089 (J.T.W.), U01 MH64092 (A.M.A.), U01 MH64003 (B.B., S.N.C.), U01 MH63747 (P.C.K.), U01 MH64107 (J.M.), and U01 MH64088 (J.P.).

References

1. Connolly SD, Bernstein GA. Practice parameter for the assessment and treatment of children and adolescents with anxiety disorders. *J Am Acad Child Psychiatry*. 2007; 46:267–283.
2. Kendall PC, Hudson JL, Gosch E, Flannery-Schroeder E, Suveg C. Cognitive-behavioral therapy for anxiety disordered youth: a randomized clinical trial evaluating child and family modalities. *J Consult Clin Psychol*. 2008; 76:282–297. [PubMed: 18377124]
3. Ollendick, TH.; King, NJ. Evidence-based treatments for children and adolescents: issues and controversies. In: Kendall, PC., editor. *Child and Adolescent Therapy: Cognitive-Behavioral Procedures*. New York, NY: Guilford Press; 2011. p. 499-519.
4. Silverman WK, Pina AA, Viswesvaran C. Evidence-based psychosocial treatments for phobic and anxiety disorders in children and adolescents. *J Clin Child Adolesc Psychiatry*. 2008; 3:105–130.
5. Walkup JT, Albano AM, et al. Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *N Engl J Med*. 2008; 359:2753–2766. [PubMed: 18974308]
6. Silverman, WK.; Albano, AM. *The Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions*. San Antonio, TX: Psychological Corporation; 1996.
7. Guy, W. *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: 1976. Clinical Global Impressions. Revised
8. Howard KI, Moras K, Brill PL, Mortinovich Z, Lutz W. Evaluation of Psychotherapy. *Am Psychol*. 1996; 51:1059–1064. [PubMed: 8870542]
9. Lambert MJ. Psychotherapy outcome and quality improvement: introduction to the special section on patient-focused research. *J Consult Clin Psychol*. 2001; 69:147–149. [PubMed: 11393592]
10. Research on Pediatric Psychopharmacology Anxiety Study Group. The Pediatric Anxiety Rating Scale (PARS): development and psychometric properties. *J Am Acad Child Psychiatry*. 2002; 41:1061–1069.
11. Research on Pediatric Psychopharmacology Anxiety Study Group. Fluvoxamine for the treatment of anxiety disorders in children and adolescents. *N Engl J Med*. 2001; 344:1279–1285. [PubMed: 11323729]
12. Kendall PC, Compton SN, Walkup JT, et al. Clinical characteristics of anxiety disordered youth. *J Anxiety Disord*. 2010; 24:360–365. [PubMed: 20206470]
13. Verduin TL, Kendall PC. Differential occurrence of comorbidity within childhood anxiety disorders. *J Clin Child Adolesc Psychiatry*. 2001; 32:290–295.
14. Poznanski, E.; Mokros, H. *Children's Depression Rating Scale-Revised (CDRSR)*. Los Angeles CA: WPS; 1996.
15. Scahill L, Riddle MA, McSwiggin-Hardin M, et al. Children's Yale-Brown Obsessive Compulsive Scale: reliability and validity. *J Am Acad Child Psychiatry*. 1997; 36:844–852.
16. Tolin DF, Abramowitz JS, Diefenbach GJ. Defining response in clinical trials for obsessive-compulsive disorder. *J Clin Psychiatry*. 2005; 66:1549–1557. [PubMed: 16401156]
17. Frank E, Prien RF, Jarrett RB, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder. *Arch Gen Psychiatry*. 1991; 48:851–855. [PubMed: 1929776]
18. Steele M, Jensen PS, Quinn DMP. Remission versus response as the goal of therapy in ADHD: a new standard for the field? *Clin Ther*. 2006; 28:1892–1908. [PubMed: 17213010]
19. Ginsburg GS, Kendall PC, Sakolsky D, et al. Remission after acute treatment in children and adolescents with anxiety disorders: findings from the CAMS. *J Consult Clin Psychol*. 2011; 79:806–813. [PubMed: 22122292]
20. Hudson JL, Rapee RM, Deveney C, Schniering CA, Lyneham HJ, Bovopoulos N. Cognitive-behavioral treatment versus an active control for children and adolescents with anxiety disorders: a randomized trial. *J Am Acad Child Adolesc Psychiatry*. 2009; 48:533–544. [PubMed: 19318990]
21. Compton SN, Walkup JT, Albano AM, et al. Child/Adolescent Anxiety Multimodal Study (CAMS): rationale, design, and methods. *Child Adolesc Psychiatry Ment Health*. 2010; 4:1–15. [PubMed: 20051130]
22. Wagner KD, Berard R, Stein MB, et al. A multicenter, randomized, double-blind, placebo-controlled trial of paroxetine in children and adolescents with social anxiety disorder. *Arch Gen Psychiatry*. 2004; 61:1153–116. [PubMed: 15520363]

23. Wood JJ, Piacentini JC, Lindsey Bergman R, McCracken JT, Barrios V. Concurrent validity of the anxiety disorders section of the Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions. *J Clin Child Adolesc Psychiatry*. 2002; 40:335–342.
24. Silverman WK, Saavedra LM, Pina AA. Test-retest reliability of anxiety symptoms and diagnoses with the Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions. *J Am Acad Child Psychiatry*. 2001; 40:937–944.
25. Langley AK, Bergman RL, McCracken J, Piacentini JC. Impairment in childhood anxiety disorders: preliminary examination of the Child Anxiety Impact Scale–Parent Version. *J Child Adolesc Psychopathol*. 2004; 14:105–114.
26. Storch, Ea; Lewin, AB.; De Nadai, AS.; Murphy, TK. Defining treatment response and remission in obsessive-compulsive disorder: a signal detection analysis of the Children's Yale–Brown Obsessive Compulsive Scale. *J Am Acad Child Psychiatry*. 2010; 49:708–717.
27. Swets, JA.; Pickett, RM. *Evaluation of Diagnostic Systems: Methods from Signal Detection Theory*. New York, NY: Academic Press; 1982.
28. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988; 240:1285–1293. [PubMed: 3287615]
29. Swets JA, Dawes RM, Monahan J. Psychological science can improve diagnostic decisions. *Psychol Sci Public Interest*. 2000; 1:1–26.
30. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950; 31:32–35. [PubMed: 15405679]
31. Kraemer, HC. *Evaluating Medical Tests: Objective and Quantitative Guidelines*. Thousand Oaks, CA: Sage Publications; 1992.
32. Kraemer HC, Periyakoil VS, Noda A. K coefficients in medical research. *Stat Med*. 2002; 21:2109–2129. [PubMed: 12111890]
33. Ginsburg GS, Keeton CP, Drazdowski TK, Riddle MA. The utility of clinician ratings of anxiety using the Pediatric Anxiety Rating Scale (PARS). *Child and Youth Care Forum*. 2010; 40:93–105.
34. Ginsburg GS. The Child Anxiety Prevention Study: intervention model and primary outcomes. *J Consult Clin Psychol*. 2009; 77:580–587. [PubMed: 19485597]
35. Ginsburg, GS.; Riddle, MA. *Pediatric Anxiety Rating Scale: revisions and normative data*. Groton, CT: Grant from Pfizer Pharmaceuticals; 2005.
36. Southam-Gerow MA, Weisz JR, Kendall PC. Youth with anxiety disorders in research and service clinics: examining client differences and similarities. *J Clin Child Adolesc Psychology*. 2003:375–385.
37. Leucht S, Davis JM, Engel RR, Kane JM, Wagenpfeil S. Defining “response” in antipsychotic drug trials: recommendations for the use of scale-derived cut-offs. *Neuropsychopharmacology*. 2007; 32:1903–1910. [PubMed: 17287825]
38. Owens EB, Hinshaw SP, Kraemer HC, et al. Which treatment for whom for ADHD? Moderators of treatment response in the MTA. *J Consult Clin Psychol*. 2003; 71:540–552. [PubMed: 12795577]
39. Chen WJ, Faraone SV, Biederman J, Tsuang MT. Diagnostic accuracy of the Child Behavior Checklist scales for attention-deficit hyperactivity disorder: a receiver-operating characteristic analysis. *J Consult Clin Psychol*. 1994; 62:1017–1025. [PubMed: 7806710]
40. Yalom ID, Lieberman MA. A study of encounter group casualties. *Arch Gen Psychiatry*. 1971; 25:16–30. [PubMed: 5096846]
41. Hannan C, Lambert MJ, Harmon C, et al. A lab test and algorithms for identifying clients at risk for treatment failure. *J Clin Psychol*. 2005; 61:155–163. [PubMed: 15609357]

FIGURE 1.

Quality index of efficiency, $\kappa(0.5)$, for Pediatric Anxiety Rating Scale (PARS) percent reduction cut-offs predicting treatment response (per Clinical Global Impression [CGI]–Improvement) and remission (per CGI-Severity and Anxiety Disorders Interview Schedule for *DSM-IV*–Child and Parent Versions [ADIS-IV-C/P]) for the total sample ($N = 438$).

TABLE 1

Prediction of Treatment Response (per Clinical Global Impression [CGI]–Improvement) at Varying Pediatric Anxiety Rating Scale (PARS) Percent Reduction Cut-offs (N = 438)

PARS Reduction (%)	Sensitivity ^a	Specificity ^b	Positive Predictive Value ^c	Negative Predictive Value ^d	Efficiency ^e	$\kappa(0.5)^f$	$\kappa(0)^g$	$\kappa(1)^h$	j^i
5	0.99	0.26	0.71	0.95	0.74	0.30	0.93	0.18	0.25
10	0.99	0.36	0.74	0.95	0.77	0.41	0.92	0.26	0.35
15	0.99	0.43	0.76	0.94	0.79	0.47	0.91	0.32	0.41
20	0.98	0.53	0.79	0.94	0.82	0.57	0.91	0.42	0.51
25	0.97	0.64	0.83	0.93	0.85	0.66	0.88	0.52	0.61
30	0.96	0.72	0.86	0.91	0.88	0.72	0.86	0.61	0.68
35	0.95	0.79	0.89	0.89	0.89	0.75	0.83	0.69	0.73
40	0.90	0.84	0.91	0.82	0.88	0.74	0.72	0.75	0.74
45	0.85	0.91	0.95	0.76	0.87	0.72	0.63	0.84	0.75
50	0.81	0.94	0.96	0.73	0.85	0.70	0.58	0.89	0.75
55	0.75	0.97	0.98	0.68	0.83	0.65	0.50	0.95	0.72
60	0.67	0.97	0.98	0.62	0.78	0.57	0.41	0.94	0.64
65	0.60	0.98	0.98	0.57	0.74	0.50	0.34	0.95	0.58
70	0.50	0.98	0.98	0.52	0.67	0.40	0.25	0.94	0.48
75	0.41	0.99	0.98	0.48	0.61	0.32	0.19	0.95	0.40
80	0.34	0.99	0.99	0.45	0.57	0.26	0.15	0.97	0.33
85	0.28	1.00	1.00	0.43	0.53	0.21	0.12	1.00	0.28
90	0.23	1.00	1.00	0.41	0.50	0.17	0.09	1.00	0.23
95	0.18	1.00	1.00	0.40	0.47	0.13	0.07	1.00	0.18

Note: Statistics for the cut-off with the highest quality of efficiency are shown in boldface type.

^aProbability of exceeding the PARS cut-off among children meeting criterion for treatment response (CGI-Improvement of 1 or 2)

^bProbability of not exceeding the PARS cut-off among those patients not meeting criterion for treatment response (CGI-Improvement of 1 or 2)

^cProbability of meeting criterion for treatment response (CGI-Improvement of 1 or 2) for those children who exceed the PARS cut-off.

^dProbability of not meeting criterion for treatment response (CGI-Improvement of 1 or 2) for those children who do not exceed the PARS cut-off.

^eProbability that the PARS cut-off and the dichotomous CGI-Improvement rating (CGI-Improvement of 1–2 vs. CGI-Improvement of 3–7) agree.

- f_i Weighted κ statistic measuring quality of efficiency.
- g_i Weighted κ statistic measuring quality of specificity.
- h_i Weighted κ statistic measuring quality of sensitivity.
- i_i Youden's J measuring optimal tradeoff of sensitivity and specificity at any single cut-off point.

TABLE 2

Prediction of Remission (per Anxiety Disorders Interview Schedule for DSM-IV–Child and Parent Versions [ADIS-IV-C/P]) at Varying Pediatric Anxiety Rating Scale (PARS) Percent Reduction Cut-offs (N = 438)

PARS Reduction (%)	Sensitivity ^a	Specificity ^b	Positive Predictive Value ^c	Negative Predictive Value ^d	Efficiency ^e	$\kappa(0.5)^f$	$\kappa(0)^g$	$\kappa(1)^h$	j^i
5	1.00	0.21	0.59	1.00	0.63	0.12	0.22	1.00	0.21
10	1.00	0.29	0.62	1.00	0.67	0.18	0.30	1.00	0.29
15	1.00	0.34	0.63	0.99	0.69	0.21	0.35	0.97	0.33
20	0.99	0.42	0.66	0.98	0.72	0.27	0.42	0.96	0.41
25	0.99	0.50	0.69	0.97	0.76	0.34	0.50	0.95	0.49
30	0.98	0.58	0.72	0.96	0.79	0.41	0.57	0.92	0.55
35	0.97	0.64	0.75	0.96	0.82	0.48	0.63	0.92	0.61
40	0.92	0.68	0.77	0.89	0.81	0.50	0.61	0.79	0.61
45	0.89	0.78	0.82	0.87	0.84	0.62	0.68	0.75	0.67
50	0.87	0.83	0.86	0.85	0.85	0.69	0.71	0.72	0.71
55	0.82	0.88	0.89	0.81	0.85	0.76	0.70	0.65	0.70
60	0.76	0.92	0.91	0.77	0.83	0.81	0.67	0.56	0.67
65	0.68	0.93	0.91	0.72	0.80	0.82	0.60	0.47	0.61
70	0.57	0.95	0.92	0.66	0.75	0.84	0.50	0.36	0.52
75	0.48	0.97	0.94	0.62	0.71	0.87	0.43	0.28	0.44
80	0.40	0.98	0.96	0.59	0.67	0.91	0.37	0.23	0.38
85	0.33	0.99	0.96	0.56	0.64	0.92	0.30	0.18	0.31
90	0.27	10.00	0.98	0.55	0.61	0.97	0.25	0.15	0.27
95	0.22	10.00	1.00	0.53	0.58	1.00	0.21	0.12	0.22

Note: Statistics for the cut-off with the highest quality of efficiency are shown in boldface type.

^aProbability of exceeding the PARS cut-off among those children meeting ADIS-IV-C/P criterion for clinical remission.

^bProbability of not exceeding the PARS cut-off among those children not meeting ADIS-IV-C/P criterion for clinical remission.

^cProbability of meeting ADIS-IV-C/P criterion for clinical remission for those children who exceed the PARS cut-off.

^dProbability of not meeting ADIS-IV-C/P criterion for clinical remission for those children who do not exceed the PARS cut-off.

^eProbability that the PARS cut-off and the ADIS-IV-C/P agree.

- f_i Weighted κ statistic measuring quality of efficiency.
- g_i Weighted κ statistic measuring quality of specificity.
- h_i Weighted κ statistic measuring quality of sensitivity.
- i_i Youden's J measuring optimal tradeoff of sensitivity and specificity at any single cut-off point.

TABLE 3

Prediction of Remission (per Anxiety Disorders Interview Schedule for DSM-IV–Child and Parent Versions [ADIS-IV-C/P]) at Varying Pediatric Anxiety Rating Scale (PARS) Absolute Cut-off Scores (N = 438)

PARS Raw Score	Sensitivity ^a	Specificity ^b	Positive Predictive Value ^c	Negative Predictive Value ^d	Efficiency ^e	$\kappa(0.5)^f$	$\kappa(1)^h$	J ⁱ
5	0.59	0.98	0.97	0.68	0.77	0.55	0.40	0.92
6	0.66	0.95	0.93	0.71	0.80	0.60	0.46	0.86
7	0.78	0.93	0.92	0.79	0.85	0.70	0.60	0.83
8	0.84	0.90	0.91	0.83	0.87	0.74	0.69	0.80
9	0.91	0.86	0.88	0.89	0.89	0.77	0.79	0.75
10	0.94	0.82	0.86	0.93	0.89	0.77	0.87	0.70
11	0.97	0.78	0.83	0.95	0.88	0.75	0.91	0.64
12	0.98	0.72	0.80	0.97	0.86	0.71	0.95	0.57
13	0.99	0.63	0.75	0.98	0.82	0.64	0.96	0.48
14	0.99	0.56	0.72	0.98	0.79	0.56	0.97	0.40
15	1.00	0.46	0.68	1.00	0.75	0.47	1.00	0.31

Note: Statistics for the cut-off with the highest quality of efficiency are shown in boldface type.

^aProbability of exceeding the PARS cut-off among those children meeting ADIS-IV-C/P criterion for clinical remission.

^bProbability of not exceeding the PARS cut-off among those children not meeting ADIS-IV-C/P criterion for clinical remission.

^cProbability of meeting ADIS-IV-C/P criterion for clinical remission for those children who exceed the PARS cut-off.

^dProbability of not meeting ADIS-IV-C/P criterion for clinical remission for those children who do not exceed the PARS cut-off.

^eProbability that the PARS cut-off and the ADIS-IV-C/P agree.

^fWeighted κ statistic measuring quality of efficiency.

^gWeighted κ statistic measuring quality of specificity.

^hWeighted κ statistic measuring quality of sensitivity.

ⁱYouden's J measuring optimal tradeoff of sensitivity and specificity at any single cut-off point.