

Defining Treatment Response in Pediatric Tic Disorders: A Signal Detection Analysis of the Yale Global Tic Severity Scale

Eric A. Storch, Ph.D.,^{1,2,3} Alessandro S. De Nadai, M.A.,² Adam B. Lewin, Ph.D.,¹ Joseph F. McGuire, B.A.,²
Anna M. Jones, B.S.,¹ P. Jane Mutch, Ph.D.,¹ R. Doug Shytle, Ph.D.,³ and Tanya K. Murphy, M.D.^{1,3}

Abstract

Objective: To examine the optimal Yale Global Tic Severity Scale (YGTSS) percent reduction and raw cutoffs for predicting treatment response among children and adolescents with tic disorders.

Method: Youth with a tic disorder ($N=108$; range = 5–17 years) participated in several clinical trials involving varied medications or psychosocial treatment, or received naturalistic care. Assessments were conducted before and after treatment and included the YGTSS and response status on the Clinical Global Impressions-Improvement Scale (CGI-I).

Results: A 35% reduction on the YGTSS total tic severity score or a YGTSS raw total tic severity score change of 6 or 7 points were the best indicators of clinical treatment response in youth with tic disorders.

Conclusions: A YGTSS total tic severity score reduction of 35% or a raw total tic severity score change of 6 or 7 appears optimal for determining treatment response. A consistent definition of treatment response on the YGTSS may facilitate cross-study comparability. Practitioners can use these values for treatment planning decisions (e.g., change medications, etc.).

Introduction

VARIED PHARMACOLOGICAL AGENTS and a psychosocial behavioral intervention have shown efficacy in the treatment of youth with tic disorders (Singer 2010). Across trials, treatment response has typically been measured through several methods, with the Yale Global Tic Severity Scale (YGTSS) total tic severity score (Leckman et al. 1989) being among the most common (McCracken et al. 2008; Murphy et al. 2009; Scahill et al. 2001; Scahill et al. 2003). The YGTSS provides a multidimensional assessment of tic symptom severity, assessing the symptom number, frequency, intensity, complexity, and interference of both motor and phonic tics over the past 7–10 days.

The YGTSS provides a rich source of data regarding tic symptom severity. However, quantitatively derived criteria for defining treatment response when using the YGTSS are lacking. Specifically, there are few data on the extent to which a certain degree of symptom reduction is associated with clinical impressions of response. This absence of established criteria in defining response can complicate interpretation of the YGTSS in both research and clinical settings. In research settings, response status has generally been operationalized across clinical trials as percent reductions on

continuous rating scales (e.g., the YGTSS) and dichotomous ratings of improvement on the Clinical Global Impressions-Improvement Scale (CGI-I) (Guy 1976). While the CGI-I provides estimates of how many subjects improved with treatment, it is not clear how this corresponds with actual reductions in tic symptom severity. Conversely, although a percent reduction in the YGTSS provides an estimate of average symptom reduction across subjects, it is an imprecise measure of how many subjects achieved a clinically meaningful response. Thus, there is a need to understand the degree of symptom alleviation associated with treatment response, which will facilitate interpretation of and comparison across clinical trials. Difficulties in defining treatment response for pediatric tic disorders can also present difficulty in the clinical setting. For example, clinicians may define clinical improvement at varying YGTSS percent reductions, where one criterion may be more accurate than others. This lack of set criteria for defining response may impact treatment planning decisions about when to revise a pharmacological or psychosocial treatment regimen e.g., if the clinician should continue with the current treatment course (if a patient is responding) or augment/switch therapies with/to another modality (if a patient is not responding). Quantitatively justified guidelines are needed for determining whether symptom reduction

¹Department of Pediatrics, University of South Florida, Tampa FL.

²Department of Psychology, University of South Florida, Tampa FL.

³Department of Psychiatry and Behavioral Sciences, University of South Florida, Tampa FL.

This research was supported, in part, by grants from the Tourette Syndrome Association to EAS, a grant from Layton Bioscience, Inc., to RDS, and grants from the National Institute of Mental Health (R01 MH063914 and K23 MH01739), National Alliance for Research on Schizophrenia and Depression, and Bristol-Myers Squibb to TKM.

in youth with tics is clinically meaningful, so as to guide such clinical decision making.

While no data on defining clinical response have been published among youth with tics, three studies have been reported in patients with a related condition, obsessive-compulsive disorder. In children, Storch et al. found that a 25% reduction on the Children's Yale-Brown Obsessive-Compulsive Scale (CY-BOCS) was maximally efficient for defining treatment response while a 45–50% reduction optimally predicted symptom remission (Storch et al. 2010). In adults, Tolin et al. found that a Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) reduction of 30% optimally predicted treatment response while Lewin et al. found a Y-BOCS reduction of 35% was optimal for predicting treatment response (Lewin et al. 2010; Tolin et al. 2005). Such established cutoffs can guide interpretations of past and future clinical trials, have direct use in decision-making for patient care, and would be desirable data to have available for research and clinical practice with pediatric tic disorders.

In sum, there are several clinical and academic implications for achieving an empirically derived threshold for treatment response. First, having an empirically determined threshold for clinical response will inform clinicians of when a certain treatment response has (or has not) been achieved. This information may have utility for treatment planning, for example, should medications be switched, maintained at a stable level, or augmented with adjunctive treatment. Second, setting a constant metric for determining a clinical threshold associated with treatment response will facilitate interpretation of and comparison across clinical trials. With these implications in mind, the present study examined YGTSS percent reductions and raw difference scores that correspond with treatment response.

Method

Participants

Participants were 108 youth (18 females) who ranged in age from 5–17 years ($M=10.87$ years, $SD=2.64$ years) and were diagnosed with Tourette syndrome, chronic motor disorder or vocal tic disorder ($n=100$ for both), transient tic disorder ($n=3$) or tic disorder not otherwise specified (NOS) ($n=5$). Youth involved in this study provided written assent (with parental consent) to participate in one of several clinical studies for children with tic disorders or in a prospective study of pediatric tic and obsessive-compulsive disorder (OCD) course. Those included in the prospective study received naturalistic care with medication changes being carefully noted at each study visit (occurring about 6 weeks apart) and YGTSS/CGI-I being administered by trained raters at each visit as part of the study design. The clinical studies included a double-blinded, placebo-controlled clinical trial of mecamlamine ($n=54$; Silver et al. 2001), an open trial of aripiprazole ($n=16$; Murphy et al. 2009), a double-blinded, placebo-controlled clinical trial of cefdinir ($n=7$; Murphy, unpublished data), an open trial of psychosocial intervention involving habit reversal training and cognitive-behavioral therapy ($n=6$; Storch unpublished data), and finally, naturalistic treatment of youth with a tic disorder within the last author's prospective tic and OCD study ($n=25$). Participants were predominantly Caucasian (93%); 4% of the sample was Hispanic, 2% African American, and 1% Asian. Co-morbidity was common among the participants with the most frequent conditions including attention-deficit/hyperactivity disorder (ADHD; 58%), OCD (44%) and oppositional defiant disorder (ODD; 38%). In regards to tic symptom severity at baseline, par-

ticipants had an average total tic severity score of 25.13 ($SD=7.92$; range=11–45) with an average impairment rating of 27.35 ($SD=10.60$; range=0–50).

Treatment

As noted, youth received a variety of interventions both in clinical trials and through naturalistic care in the context of a prospective study on tic and OCD course. It is important to note that the mechanism of change (i.e., the intervention type) is not a factor as long as the YGTSS cutoff reflects what is clinically viewed as treatment response. Mecamlamine, a non-selective antagonist of the nicotinic acetylcholine receptors, was well-tolerated in doses up to 7.5 mg/day but not efficacious relative to placebo (Silver et al. 2001). Aripiprazole, an atypical antipsychotic, was well-tolerated in children with tics and associated with significant tic reduction (Murphy et al. 2009). Cefdinir is an antibiotic; its effects on tics with possible immune-related etiology are being investigated. The psychosocial intervention included habit reversal training for tics and cognitive-behavioral skills to address psychosocial problems that often accompany tic disorders (e.g., peer victimization, poor academic functioning, etc.) delivered over 12 sessions. For the purposes of this study, naturalistic care within the last author's prospective study involved standard clinical practice for the pharmacotherapeutic treatment of tics, which included office visits approximately every six weeks. Medications prescribed and corresponding dosages were based on practice guidelines for the treatment of pediatric tics (e.g., Scahill et al. 2006; Singer 2010) together with the provider's clinical judgment. During office visits, the child met with the psychiatrist for ≥ 30 minutes, during which time the child's clinical presentation, medication response, and any medication-related side effects were discussed in a supportive clinical environment.

Overall, 68 participants (63%) were receiving one or more active psychiatric medications at therapeutic doses that have been shown to influence tics as their primary treatment (Scahill et al. 2006; Singer 2010). This included 27 (25%) participants on nicotinic acetylcholine antagonists (e.g., mecamlamine), 20 (19%) on an antipsychotic without concurrent psychosocial therapy (e.g., aripiprazole, risperidone), and 5 (5%) on an alpha agonist (e.g., clonidine, guanfacine). Sixteen participants (15%) were receiving multiple tic influencing medications which consisted of at least one antipsychotic medication. Five participants (5%) received cefdinir. Six participants (6%) received a psychosocial intervention, two of which were also taking a stable dose of aripiprazole. The remaining 29 participants (27%) received placebo as part of double-blind placebo-controlled trials for either mecamlamine ($n=27$) or cefdinir ($n=2$); patients on placebo were included in analyses.

Measures

The following measures were administered by trained independent evaluators specific to each study and are briefly described. A structured diagnostic interview to confirm primary and comorbid diagnoses was used. Depending on the study, this consisted of one of the following diagnostic interviews: the Mini-International Neuropsychiatric Interview (Sheehan et al. 1998), Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime version (K-SADS; Kaufman et al. 1997), or the Anxiety Disorders Interview Schedule for DSM-IV: Parent Version with a supplementary tic module (Silverman and Albano 1996). As noted above, the YGTSS is a clinician-administered interview that evaluates motor and phonic symptom number, frequency, intensity,

complexity, and interference, as well as global impairment from tics (Leckman et al. 1989). The CGI-I (Guy 1976) is a clinician-rated scale where clinical improvement is rated from 1 (“very much improved”) to 7 (“very much worse”). A rating of “much improved” or “very much improved” was used to designate treatment response, consistent with published randomized controlled trials (e.g., Piacentini et al. 2010). While there are advantages to using the CGI-I as a measure of overall treatment response (e.g., simplicity, ability to provide an assessment of gestalt improvement), the CGI-I has several inherent limitations including: 1) problems with inter-reliability and rater-experience (e.g., raters may base ratings on past experiences); 2) difficulty accounting for inter-patient variability in how symptom reduction is perceived/experienced (e.g., two patients who experience a 25% reduction may show differing levels of overall improvement); and 3) the simple format of the CGI-I may limit its value for the week-to-week tracking of symptoms outside of a clinical trial (i.e., a lack of symptom detail). Despite such issues, the CGI-I is well-regarded and widely-used across mental health treatment outcome trials as a global measure of treatment response.

Procedures

Institutional review board granted permission for all research studies, and written parent consent and child assent was obtained prior to involvement in the respective study, as well as to allow data to be used for other research in tic disorders. Prior to the initial assessment, a Diagnostic and Statistical Manual for Mental Disorders (APA 2000) diagnosis of a tic disorder (e.g., Tourette syndrome, chronic motor or vocal tic disorder, transient tic disorder or tic disorder NOS) was established using both semi-structured diagnostic (i.e., K-SADS) and clinical interviews. Subsequently, comprehensive assessments conducted by independent evaluators occurred at baseline and post-treatment (as well as other time-points not germane to the present investigation).

At the baseline and post-treatment assessments, a clinician not involved in treatment administered the YGTSS to parents and children to evaluate tic symptom severity. After the baseline assessment, participants began their respective study intervention or naturalistic care consisting of a tic influencing medication (e.g., risperidone, guanfacine, etc.). The same clinician later made CGI-I ratings based on their impressions of treatment response using all available information (e.g., YGTSS, patient response).

Analytic plan

We aimed to assess the performance of various YGTSS total tic severity score percent reduction and raw total tic difference cutoff scores in detecting treatment response (as defined by the CGI-I). To achieve these ends, we used receiver operating characteristic (ROC) methods derived from signal detection theory (Swets and Pickett 1982). ROC analyses focus on ratios of diagnostic true positive, false positive, true negative and false negative results. In assessing the performance of YGTSS total tic severity score percent reduction cutoffs in detecting clinical response, we examined percent reductions from 5–70% at 5% intervals, following precedent from previous research studies that employed this methodology (Storch et al. 2010; Tolin et al. 2005). To examine raw difference YGTSS total tic cutoff scores, we examined a range of possible YGTSS total tic severity scores that would be likely to accurately detect clinical response. In the context of this study, each YGTSS total tic severity cutoff/score is treated as a “rater.” Accordingly, our goal is to identify which

cutoff (or “rater”) has the best psychometric properties in detecting clinical response.

In assessing the performance of the YGTSS in detecting clinical response, we used the ROC statistics of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and efficiency (also known as accuracy). Sensitivity is the probability of exceeding the test cutoff among those meeting criteria for clinical response (i.e., the probability that the YGTSS total tic severity score cutoff captures a true responder). Specificity is the probability of not exceeding the test cutoff among those who do not meet criteria for clinical response (i.e., the probability that the YGTSS total tic severity score cutoff captures a true non-responder). Positive predictive value is the probability of displaying clinical response should a patient exceed the YGTSS total tic severity score cutoff, and negative predictive value is the probability of not displaying clinical response should a patient not exceed the YGTSS total tic severity score. Efficiency (also known as accuracy) is the probability that the YGTSS total tic severity cutoff score and the CGI-I agree in their assessment of clinical response.

Given that many gold standard outcomes still retain at least minimal measurement error (in this case, the CGI-I), Kraemer and colleagues (Kraemer 1992; Kraemer et al. 2002) have proposed a series of weighted Kappa statistics to correct for such error, and employ these to assess for the quality of ROC statistics under the definition of Quality Receiver Operating Characteristic (QROC) Methods. Of particular importance to answering the present research questions are the K(0), K(.5), and K(1) statistics, which measure the quality of specificity, quality of efficiency, and quality of sensitivity, respectively. In interpreting these statistics, a value of 0.00 indicates these properties cannot be differentiated from chance, and a value of 1.00 indicates perfect assessment by a cutoff with respect to each ROC property. Sensitivity is often valued in screening purposes (e.g., when one wants to ensure that all true positive results are captured by a measure and that no true positive results are left behind). Conversely, specificity is often valued when there is a need for a definitive diagnosis (e.g., when one needs to employ a costly intervention and wishes to be certain that true negatives are being excluded and not being kept for consideration). Given our interest in making the most accurate judgment at any given time, we placed the highest value on maximizing efficiency (and thus the value of the K(.5) statistic, which weights false positive and false negative results equally). Lastly, positive and negative predictive value function much as one would use a diagnostic instrument in the real world (e.g., given that one is considered a responder by this measure, what is the probability that he/she is actually a responder?). Given this inference, the false positive rate can be calculated as $1-PPV$, and the false negative rate is $1-NPV$.

Results

Determining treatment response based on YGTSS percentage reduction

Appropriate ROC and QROC statistics for assessing the performance of YGTSS total tic severity score percent reduction cutoffs in detecting clinical response are presented in Table 1. Optimal efficiency in detecting clinical response (as measured by the K(.5) statistic) was found at the 35% reduction cutoff, with 79% of treatment responders being accurately detected by this cutoff; however, sensitivity was somewhat low at .70, suggesting that using YGTSS total tic severity score percent reduction cutoffs to detect clinical response may be somewhat insensitive and miss

TABLE 1. SIGNAL DETECTION ANALYSES OF THE PREDICTION OF CLINICAL RESPONSE (IMPROVED OR VERY MUCH IMPROVED ON THE CLINICAL GLOBAL IMPRESSIONS-IMPROVEMENT SCALE) AT VARYING YALE GLOBAL TIC SEVERITY SCALE (YGTSS)-TOTAL PERCENT REDUCTION CUTOFF SCORES

YGTSS Reduction (%)	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Efficiency	K(0.0)	K(1.0)	K(0.5)
≥ 5	0.87	0.52	0.64	0.80	0.69	0.29	0.60	0.39
≥ 10	0.83	0.56	0.65	0.77	0.69	0.30	0.54	0.39
≥ 15	0.81	0.63	0.69	0.77	0.72	0.38	0.55	0.44
≥ 20	0.80	0.70	0.73	0.78	0.75	0.46	0.55	0.50
≥ 25	0.78	0.76	0.76	0.77	0.77	0.53	0.55	0.54
≥ 30	0.72	0.81	0.80	0.75	0.77	0.59	0.49	0.54
≥ 35	0.70	0.87	0.84	0.75	0.79	0.69	0.49	0.57
≥ 40	0.69	0.87	0.84	0.73	0.78	0.68	0.47	0.56
≥ 45	0.61	0.89	0.85	0.70	0.75	0.69	0.39	0.50
≥ 50	0.56	0.93	0.88	0.68	0.74	0.76	0.35	0.48
≥ 55	0.44	0.98	0.96	0.64	0.71	0.92	0.28	0.43
≥ 60	0.33	0.98	0.95	0.60	0.66	0.89	0.19	0.31
≥ 65	0.24	1.0	1.0	0.57	0.62	1.0	0.14	0.24
≥ 70	0.19	1.0	1.0	0.55	0.59	1.0	0.10	0.19

some true responders. The 30% and 40% cutoffs display similar (albeit slightly reduced) performance in ROC and QROC statistics, but with no appreciable improvement in performance in any ROC or QROC domain. The 25% cutoff also shows similar performance, with higher sensitivity at .78, but with lower specificity at .76 and a lower PPV at .76 (and thus a false positive rate of 24%).

Determining treatment response based on YGTSS raw difference scores

Appropriate ROC and QROC statistics for assessing the performance of YGTSS raw difference cutoffs in detecting clinical response are presented in Table 2. Optimal efficiency in detecting clinical response (as measured by the K(.5) statistic) was found at a YGTSS total tic severity score raw difference of 6 and 7, with 77% of treatment responders being accurately detected at this cutoff. A raw difference cutoff of 6 had a slightly better sensitivity than that of 7 (.78 vs. .74), at the expense of slightly poorer specificity (.76 vs. .80), with both cutoffs having PPV and NPV differences no larger than .02.

Visually charting the performance of the YGTSS in detecting treatment response

The K(0.5) metric in assessing the performance of varying YGTSS percent reduction and raw difference cutoffs can be found in Figures 1 and 2, respectively. The peak points on these figures correspond to the points of maximal efficiency identified in Tables 1 and 2. One point of note is that Figure 2 (corresponding to raw difference cutoffs) is somewhat flat, indicating that optimal performance is indeed at a cutoff of 6 or 7, but also that these cutoffs are not maximally differentiated in efficiency from other cutoff options; however, specificity is substantially lower at many of such comparable cutoffs.

Discussion

The YGTSS has been widely utilized as an outcome measure in treatment studies for individuals with tic disorders and has utility for practitioners in gauging treatment response to pharmacological and psychosocial interventions. To date, however, there are no

TABLE 2. SIGNAL DETECTION ANALYSES OF THE PREDICTION OF CLINICAL RESPONSE (IMPROVED OR VERY MUCH IMPROVED ON THE CLINICAL GLOBAL IMPRESSIONS-IMPROVEMENT SCALE) AT VARYING YALE GLOBAL TIC SEVERITY SCALE (YGTSS)-TOTAL RAW DIFFERENCE CUTOFF SCORES

YGTSS Cutoff	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Efficiency	K(0.0)	K(1.0)	K(0.5)
1	0.87	0.43	0.60	0.77	0.65	0.21	0.53	0.30
2	0.87	0.52	0.64	0.80	0.69	0.29	0.60	0.39
3	0.81	0.57	0.66	0.76	0.69	0.31	0.51	0.39
4	0.80	0.63	0.68	0.76	0.71	0.37	0.51	0.43
5	0.78	0.72	0.74	0.76	0.75	0.47	0.53	0.50
6	0.78	0.76	0.76	0.77	0.77	0.53	0.55	0.54
7	0.74	0.80	0.78	0.75	0.77	0.57	0.51	0.54
8	0.70	0.81	0.79	0.73	0.76	0.58	0.47	0.52
9	0.65	0.87	0.83	0.71	0.76	0.67	0.42	0.52
10	0.59	0.91	0.86	0.69	0.75	0.73	0.38	0.50
11	0.59	0.91	0.86	0.69	0.75	0.73	0.38	0.50
12	0.57	0.93	0.89	0.68	0.75	0.77	0.37	0.50
13	0.52	0.93	0.88	0.66	0.72	0.75	0.32	0.44
14	0.44	0.94	0.89	0.63	0.69	0.78	0.26	0.39

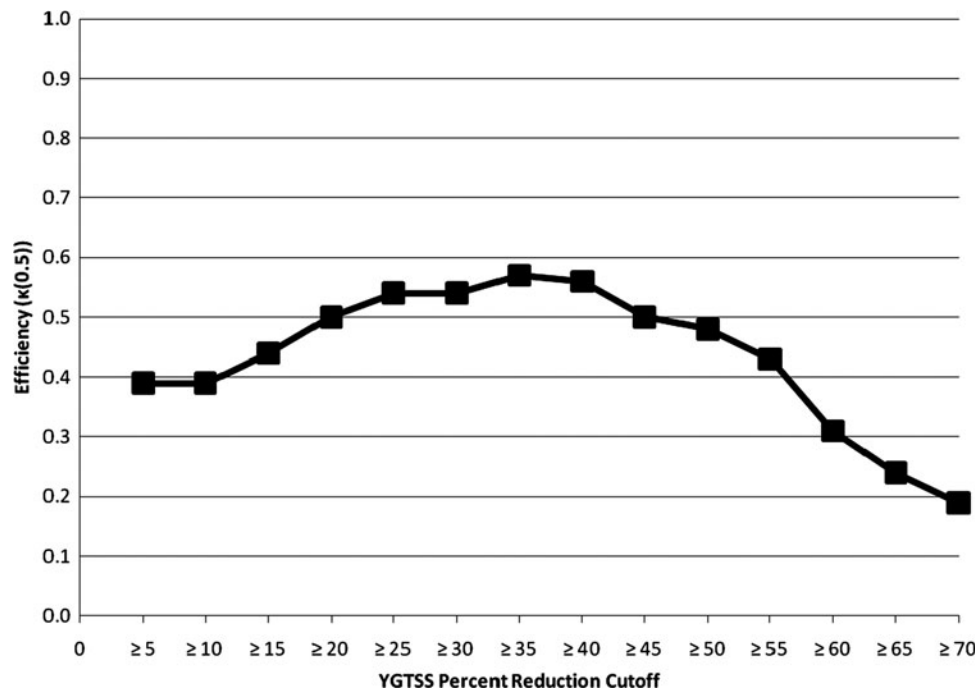


FIG. 1. Quality Index of Efficiency ($\kappa(0.5)$) for the predictive value of differing Yale Global Tic Severity Scale (YGTSS) percent reduction cutoffs relative to predicting clinical response.

empirical guidelines for determining optimal tic symptom reduction and cutoff scores that correspond with clinical improvement. Consequently, the present study sought to formally establish the optimal criteria for judging treatment response using the YGTSS.

The optimal YGTSS total tic severity score percent reduction cutoff for determining response was 35%, when balancing the costs of false positives and false negatives equally (which is the goal in clinical

trials). In a clinical setting, however, sometimes false negatives are more problematic than false positives. That is, a clinician does not want to abandon a treatment course for a patient that is actually responding. Given these scenarios, a 25% reduction could be used to determine response in clinical settings, given the slightly poorer performance on $\kappa(.5)$ (which weights false positives and negatives equally), but displays better sensitivity (and stronger $\kappa(1)$ performance), meaning that

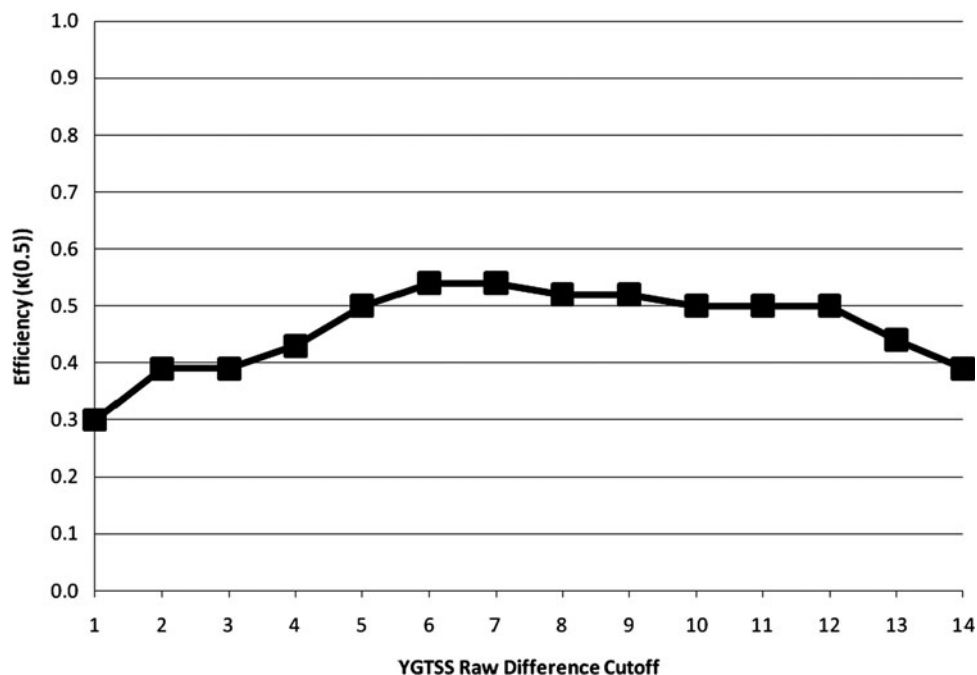


FIG. 2. Quality Index of Efficiency ($\kappa(0.5)$) for the predictive value of differing Yale Global Tic Severity Scale (YGTSS) raw difference cutoffs relative to predicting clinical response.

fewer true responders are missed by this cutoff than at the 35% cutoff. Inflexibly requiring a higher percent YGTSS total tic severity score reduction in clinical contexts may result in changing interventions prematurely. In assessing the performance of raw difference YGTSS cutoffs to denote treatment response, YGTSS total tic severity scores of 6 or 7 were considered optimal. Raw score differences are less influenced by initial symptom severity and are readily communicable (i.e., it is easy to quickly assess how many points a patient has reduced when making treatment planning decisions).

Despite offering suggested cutoff scores and percent symptom reduction targets, ROC and QROC analyses found that YGTSS total tic severity score did not perform optimally in detecting treatment response. This may be a product of the inherent heterogeneity of tic presentation, with patients differing in motor and phonic tic severity; in creating a total score, patients with predominantly one symptom type may not be measured similarly to patients with equal motor and phonic tic severity. However, simply breaking down analyses by YGTSS Motor and Phonic score cutoffs can suffer from the same problem, where a patient who has all motor tics may be measured accurately but one who has additional phonic tics will not be measured as accurately when using only one of the two component scales. For example, if two patients have a YGTSS total tic severity score of 25, one patient who has a YGTSS Phonic score of 25 (the maximum attainable score) may present differently than a patient with a YGTSS Motor score of 13 and a YGTSS Phonic score of 12. In considering these two patients in the context of treatment response, if the first patient has a YGTSS Phonic reduction of approximately 50% and goes from a score of 25 to a score of 12, this change may be qualitatively different than the same YGTSS total tic percent reduction in the second patient that could render a score of 6 on both the YGTSS Motor and Phonic subscales. Thus, although the YGTSS is adequately able to detect clinical response, these data indicate that reconceptualization of the YGTSS scoring structure may enhance its measurement sensitivity.

Several limitations of this research are noteworthy. First, participants were drawn from several different studies offering different interventions. However, the aim of the study was to identify optimal YGTSS scores and percent reductions based on clinical response. The focus of these analyses are to detect clinical change; the mechanism of change (i.e., the intervention type) is not empirically relevant unless it is hypothesized that the YGTSS detects clinical change differently depending on its cause (e.g., a patient responding to pharmacotherapy will not be as accurately detected by the YGTSS than a patient who responds to psychotherapy). Given the conventional use of the YGTSS to assess outcome across pharmacotherapy and psychotherapy studies, there is no current evidence to indicate this is the case. In fact, sample heterogeneity tends to improve generalizability and the variation in treatment approaches may better reflect how the YGTSS is used in a variety of settings. Second, we were unable to examine optimal remission cutoffs associated with the YGTSS. However, unlike some psychiatric disorders, clinically significant improvement rather than remission may be the more common and realistic target in youth with tic disorders. Third, we were unable to group our ROC analyses by age, gender, or illness characteristics (e.g., YGTSS subscales) given the current sample size; however, there is no *a priori* reasoning to expect why the YGTSS should perform differently among these populations. Fourth, the sample only included children and as such, these findings may not be generalized to adults with tics. Fifth, unfortunately data were not available to examine the degree to which CGI-I ratings and YGTSS total tic severity scores covary at different time points during treatment. Finally, the

present analysis focuses on the YGTSS total tic severity score as a predictor of treatment response. It is important to consider that patients may present with both phonic and motor tics or with either predominantly phonic or motor tics. These differing classes of patients may not be reflected in the YGTSS total tic severity score (e.g., a child with severe motor tics [no phonic tics] may have a similar YGTSS total tic severity score as a child with moderate phonic and motor tics). Nevertheless, although this inherent limitation of the YGTSS total tic scoring may impact the measurement of tic severity, it does not impact the signal detection analysis which examines change in relation to the CGI-I.

Clinical Significance

Results of this research provide a set of guidelines for judging response using the YGTSS and have implications for both researchers and practitioners. For treatment outcome research, these data provide empirically defined parameters for determining cutoffs for response, which may be of use across research trials (e.g., a child is considered a responder if they have a certain level of tic reduction). In clinical practice, YGTSS total tic severity score percent reduction and raw cutoffs identified in this research can guide treatment decisions, for example, when to change, augment, and/or conclude a course of intervention. In addition, these data allow professionals to identify the optimal cutoff based on the parameters of most importance (e.g., specificity, sensitivity, PPV, NPV, or efficiency) for the relevant use (i.e., treatment outcome research, screening, clinical practice).

Disclosures

Dr. Storch receives grant funding from the National Institute of Mental Health, NICHD, All Children's Hospital Research Foundation, Centers for Disease Control, National Alliance for Research on Schizophrenia and Affective Disorders, International OCD Foundation, Tourette Syndrome Association, Janssen Pharmaceuticals, and Foundation for Research on Prader-Willi Syndrome. He receives textbook honorarium from Springer publishers and Lawrence Erlbaum. Dr. Storch has been an educational consultant for Rogers Memorial Hospital. Dr. Mutch, Mr. De Nadai and McGuire, and Ms. Jones have no conflicts of interest or financial ties to disclose. Dr. Lewin receives grant funding from National Alliance for Research on Schizophrenia and Affective Disorders and International OCD Foundation. Dr. Shytle has been a consultant for Pfizer, Yaupon Therapeutics, AstraZeneca, and Natura Therapeutics. Dr. Murphy has received research support from National Institute of Mental Health; Forest Laboratories; Janssen Pharmaceuticals; Endo; Obsessive Compulsive Foundation; Tourette Syndrome Association; All Children's Hospital Research Foundation, Centers for Disease Control, National Alliance for Research on Schizophrenia and Affective Disorders, Dr. Murphy is on the Medical Advisory Board for Tourette Syndrome Association. She receives textbook honorarium from Lawrence Erlbaum.

Acknowledgment

We acknowledge the contributions of Archie Silver, M.D. to this research.

References

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders. Washington, D.C.: American Psychiatric Association, 2000.

- Guy W: Clinical Global Impressions. In: ECDEU Assessment Manual for Psychopharmacology (Vol. Revised DHEW Pub. (ADM)). Rockville, MD: National Institute for Mental Health, pp. 218–222, 1976.
- Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N: Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version: Initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 36:980–988, 1997.
- Kraemer HC. Evaluating Medical Tests: Objective and Quantitative Guidelines. Thousand Oaks, CA: Sage Publications, 1992.
- Kraemer HC, Periyakoil VS, Noda A: Kappa coefficients in medical research. *Stat Med*. 21:2109–2129, 2002.
- Leckman JF, Riddle MA, Hardin MT, Ort SI, Swartz KL, Stevenson J, Cohen DJ: The Yale Global Tic Severity Scale: initial testing of a clinician-rated scale of tic severity. *J Am Acad Child Adolesc Psychiatry* 28:566–573, 1989.
- Lewin AB, De Nadai AS, Park J, Goodman WK, Murphy TK, Storch EA: Refining clinical judgment of treatment outcome in obsessive-compulsive disorder. *Psychiatry Res* 185:394–401, 2011.
- McCracken JT, Suddath R, Chang S, Thakur S, Piacentini J: Effectiveness and tolerability of open label olanzapine in children and adolescents with Tourette syndrome. *J Child Adolesc Psychopharmacol* 18:501–508, 2008.
- Murphy TK, Kurlan R, Leckman J: The immunobiology of Tourette's disorder, pediatric autoimmune neuropsychiatric disorders associated with Streptococcus, and related disorders: A way forward. *J Child Adolesc Psychopharmacol* 20:317–331, 2010.
- Murphy TK, Mutch PJ, Reid JM, Edge PJ, Storch EA, Bengtson M, Yang M: Open label aripiprazole in the treatment of youth with tic disorders. *J Child Adolesc Psychopharmacol* 19:441–447, 2009.
- Piacentini J, Woods DW, Scahill L, Wilhelm S, Peterson AL, Chang S, Ginsburg GS, Deckersbach T, Dziura J, Levi-Pearl S, Walkup JT: Behavior therapy for children with Tourette disorder: A randomized controlled trial. *JAMA* 303: 1929–1937, 2010.
- Scahill L, Chappell PB, Kim YS, Schultz RT, Katsovich L, Shepherd E, Arnsten AF, Cohen DJ, Leckman JF: A placebo-controlled study of guanfacine in the treatment of children with tic disorders and attention deficit hyperactivity disorder. *Am J Psychiatry* 158:1067–1074, 2001.
- Scahill L, Erenberg G, Berlin CM, Jr., Budman C, Coffey BJ, Janovic J, Walkup J. Contemporary assessment and pharmacotherapy of Tourette syndrome. *NeuroRx*, 3:192–206, 2006.
- Scahill L, Leckman JF, Schultz RT, Katsovich L, Peterson BS: A placebo-controlled trial of risperidone in Tourette syndrome. *Neurology* 60:1130–1135, 2003.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC: The Mini-International Neuropsychiatric Interview: The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 59 Suppl 20:22–33, 1998.
- Silver AA, Shytle RD, Sheehan KH, Sheehan DV, Ramos A, Sanberg PR: Multicenter, double-blind, placebo-controlled study of mecamylamine monotherapy for Tourette's disorder. *J Am Acad Child Adolesc Psychiatry* 40:1103–1110, 2001.
- Silverman WK, Albano AM. The Anxiety Disorders Interview Schedule for DSM-IV-Child and Parent Versions. San Antonio, TX: Graywinds Publications, 1996.
- Singer HS: Treatment of tics and Tourette syndrome. *Curr Treat Options Neurol* 12:539–561, 2010.
- 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 Adolesc Psychiatry* 49:708–717, 2010.
- Swets JA, Pickett RM. Evaluation of Diagnostic Systems: Methods from Signal Detection Theory. New York, NY: Academic Press, 1982.
- Tolin DF, Abramowitz JS, Diefenbach GJ: Defining response in clinical trials for obsessive-compulsive disorder: a signal detection analysis of the Yale-Brown obsessive compulsive scale. *J Clin Psychiatry* 66:1549–1557, 2005.

Address correspondence to:

Eric A. Storch, Ph.D.

Department of Pediatrics

Rothman Center for Neuropsychiatry

University of South Florida

800 6th Street South 4th Floor

St. Petersburg, FL 33701

E-mail: estorch@health.usf.edu

