

Using a Brief Parent-Report Measure to Track Outcomes for Children and Teens with Internalizing Disorders

Hayley S. Kamin · Alyssa E. McCarthy ·
Madelaine R. Abel · Michael S. Jellinek ·
Lee Baer · J. Michael Murphy

Published online: 5 December 2014
© Springer Science+Business Media New York 2014

Abstract The Pediatric Symptom Checklist (PSC) is a widely-used, parent-completed measure of children's emotional and behavioral functioning. Previous research has shown that the PSC and its subscales are generally responsive to patient progress over the course of psychiatric treatment. In this naturalistic study, we examined the performance and utility of the five-item PSC Internalizing Subscale (PSC-IS) as an assessment of routine treatment in outpatient pediatric psychiatry. Parents and clinicians of 1,593 patients aged 17 or younger completed standardized measures at intake and three-month follow-up appointments. Comparisons between PSC-IS scores and clinician-reported diagnoses, internalizing symptoms, and overall functioning showed acceptable levels of agreement. Change scores on the PSC-IS were also larger among patients with internalizing diagnoses than those with non-

internalizing diagnoses. As a brief measure of internalizing symptoms, the PSC may be particularly useful to mental health clinicians treating youth with depression and anxiety as a quality assurance or treatment outcome measure.

Keywords Child psychiatry · Pediatrics · Depression · Anxiety disorders · Outcome assessment

Introduction

Depression and anxiety disorders are prevalent and burdensome conditions among children and adolescents [1–3]. Often comorbid with anxiety disorders, major depression is a leading cause of disability worldwide [4]. Youth-onset depression and anxiety in particular are associated with a number of adverse outcomes by young adulthood, including suicidal behavior, substance dependence, early parenthood, and educational underachievement [1, 3, 5]. As depressive and anxiety disorders in childhood and adolescence are highly predictive of more chronic forms of the disorders in adulthood [6], evaluation and treatment of these conditions merits greater attention to avoid costly and potentially long-term consequences.

Given the substantial impact of internalizing disorders on society, the search for effective interventions should be a high priority. Recent critiques have suggested that psychopharmacological interventions, particularly those using selective serotonin reuptake inhibitors to treat depression, may have much smaller effect sizes than originally reported [7]. If the effect sizes associated with current treatments are indeed smaller, then larger samples will be needed to detect the actual effects of psychiatric interventions experienced in clinical practice [8]. As at least one recent article [9] suggested, the growing availability of standardized

H. S. Kamin
Department of Psychology, University of Florida, Gainesville,
FL, USA

A. E. McCarthy
Department of Psychology, Rutgers University, New Brunswick,
NJ, USA

M. R. Abel · M. S. Jellinek · L. Baer · J. M. Murphy (✉)
Child Psychiatry Service, Division of Child and Adolescent
Psychiatry, Department of Psychiatry, Massachusetts General
Hospital, 55 Fruit Street, YAW 6A, Boston, MA 02114, USA
e-mail: mmurphy6@partners.org

M. S. Jellinek · L. Baer · J. M. Murphy
Department of Psychiatry, Harvard Medical School, Boston,
MA, USA

M. S. Jellinek
Department of Pediatrics, Harvard Medical School, Boston, MA,
USA

measures in clinic, health plan, state, and national datasets could provide a convenient and feasible way to harness the power of large, real-world samples to explore questions that involve small effect sizes. Consistent utilization of standardized measures has the advantages of collecting information as part of an evidence base for empirically supported treatments, while also providing clinicians with a more holistic picture of patient progress and enabling patients to become active participants in their treatment. The benefits achieved by incorporating patient feedback are seen on multiple levels of clinical care and practice accountability, and include enhanced treatment outcomes in a shorter number of sessions, improved accuracy of diagnosis, less likelihood of treatment failure, facilitated communication between patients and clinicians, and greater accuracy in evaluating service quality and determining allocation of funding [10–14].

Standardized measures have long been used to evaluate patient functioning and psychopathology [9], and there are several widely-used instruments designed specifically for youth internalizing disorders. However, most of these measures are lengthy or focused on a single diagnostic area (e.g., the 28-item Children's Depression Inventory (CDI) [15]; the 41-item Screen for Child Anxiety Related Emotional Disorders (SCARED) [16]), which makes it less likely that they will be routinely administered to large populations or be able to adequately assess more general internalizing problems. The nine- [17] and two-item [18] versions of the Patient Health Questionnaire are brief and increasingly used in real-world settings [19], but limited by the fact that they focus solely on depression.

The Pediatric Symptom Checklist (PSC) [20] is another brief measure that is widely-used in care settings. The PSC is a parent-/guardian-completed ('parent' for the remainder of this article) measure of psychosocial impairment in children and adolescents. Since it is routinely used as a repeated measure in psychiatric [21, 22], statewide pediatric [23], and national education samples [24], it holds promise as a measure that could be used to assess outcomes over time and across systems of care in existing, real-world samples. From 2011 to 2014 the PSC was endorsed for use in state- and national-scale evaluations by the National Quality Forum [25], a consensus-based organization tasked with evaluating and endorsing standards for measuring performance as part of a national strategy for healthcare quality. The PSC is among a set of measures that comprise a fundamental part of accountability and quality improvement programs nationwide [26]. In addition, as discussed in more detail below, the PSC meets a number of criteria described by Delgado et al. (2012) [27] as central features of outcome measures, namely, its strong evidence base as a valid measure of child psychosocial functioning, acceptability to clinicians and patients, ease of

administration and interpretation, widespread availability in multiple formats and languages, and extent of published data that allows for comparisons across settings.

The PSC provides both continuous and categorical (case vs. not case) scores for the global scale [20] and subscales in three areas: internalizing, externalizing, and attention problems [28]. Recent studies showed high rates of agreement among categorical cutoff scores indicating the absence or possible presence of internalizing problems on its five-item PSC Internalizing Subscale (PSC-IS), clinical diagnoses based on the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)* [29], and several common but longer measures for youth depression and anxiety [30, 31]. Two other studies of outpatient child psychiatry samples showed that the PSC global scale [22, 32] and subscales [22] registered significant improvements between intake and 3 [22, 32] and 6 months [22] into treatment. These studies suggest that it is possible to use the PSC to track treatment progress and outcomes by quantifying the degree of change on its three subscales while monitoring overall functioning with the total score. The brief PSC-IS might therefore be able to help providers identify internalizing problems and track changes in patients' internalizing symptoms over the course of treatment. This could be especially useful for pediatricians looking to monitor their patients' functioning before or after referral.

In a recently published article, Murphy and colleagues demonstrated the practical utility of the PSC's clinical thresholds for evaluating psychiatric care [32]. Using the 'reliable and clinically significant change' calculations proposed by Jacobson and Truax [33], the authors examined rates of improvement among youth treated in an outpatient psychiatric clinic. Jacobson and Truax's reliable change index (RCI) indicates the minimum amount of change on a measure unlikely to be attributable to statistical fluctuation. This change is considered 'clinically significant' when the patient's score also crosses the measure's risk threshold. This method of assessing outcomes allows for comparisons at the level of individual patients [34] and has been described as fundamental to patient-focused research [35] and service evaluation [34]. For clinicians, these metrics can aid treatment planning by identifying patients who appear to be improving, not experiencing significant change, or deteriorating [32, 36].

The RCI is a common outcome measure in randomized controlled trials of treatments for youth with internalizing disorders [37, 38]. These studies typically involve particular treatment protocols or selective sampling that excludes youth with comorbid disorders. We are not aware of any naturalistic studies that used the RCI and clinically significant change metrics to assess rates of improvement and deterioration among youth treated for internalizing disorders

in ordinary clinical settings. Since the PSC-IS is easily administered and commonly used in such settings, we believe it can be an appropriate tool for addressing the gap between carefully controlled research protocols and what happens in real-world clinical practices. If the PSC-IS accurately assesses rates of reliable and clinically significant change among youth receiving naturalistic treatments for internalizing problems, researchers and clinicians could use it to gain access to valuable information about treatment effectiveness in real-world settings and datasets.

In the current study, we examined the clinical utility of the PSC-IS as a screen and measure of treatment progress. Using a sample of youth who received outpatient psychiatric care, we evaluated the degree to which scores on the previously validated PSC-IS agreed with clinician-reported diagnoses, symptom severity, and change over the course of treatment. We also examined whether the PSC-IS was responsive to improvement and deterioration according to the reliable and clinically significant change thresholds.

Methods

Sample

Patients were eligible for inclusion in the study if they entered treatment and were assessed in the Massachusetts General Hospital (MGH) Outpatient Child and Adolescent Psychiatry Service between August 1, 2007 and July 31, 2013. As this study was intended to represent the wide diversity of patients seen in clinical practice, no exclusion criteria were applied; all patients under 17 years were requested to fill out the outcome forms as a part of routine care at the clinic and all patients with complete outcome forms were included in the study sample. Our first analytic sample included patients who had complete PSCs and DSM-IV Axis I diagnoses on their intake forms. Our second analytic sample included patients with completed PSCs at 3 months into treatment. Patients with primary and/or secondary Axis I diagnoses noted on their intake forms constituted the internalizing sample. The data presented here includes some of the data reported on previously studied samples in this clinic (August 1, 2007 to April 30, 2008 [21] and August 1, 2007 to July 31, 2010 [22]).

Procedures

This study examined ‘treatment as usual’ in a naturalistic population of children and adolescents seeking mental health care at a large, urban, academic hospital. Patients were treated with therapy and/or psychopharmacology based on the treating clinician’s best judgment. Over a 5-year period, researchers in the Division of Child and

Adolescent Psychiatry at MGH collected parent and clinician reports of patient functioning at intake and subsequent three-month follow-up appointments. All procedures were reviewed and approved by the hospital’s Human Research Committee.

Throughout this study’s data collection period, the department required completion of electronic Outcome Rating Forms (e-ORFs) at intake and every 3 months thereafter in order to track treatment progress [21]. The e-ORF for patients aged 17 and younger contains the parent-completed PSC [20] and the clinician-completed Brief Psychiatric Rating Scale for Children (BPRS-C) [39] and Children’s Global Assessment Scale (CGAS) [30]. The e-ORF also provides demographic information about the patient drawn from the hospital’s electronic scheduling system and requests that clinicians list the patient’s DSM-IV [29] diagnoses and treatment type. Parents complete their portion of the form in the waiting room before their child’s appointment and clinicians complete their portion during or after the appointment. Clinicians are encouraged to review the Pediatric Symptom Checklist (PSC) with the family at each assessment. The form and its administration have been described in greater detail in previous publications [20, 21, 40].

Measures

Parent-Completed Measure

The PSC is 35-item measure designed to evaluate emotional, behavioral, and social functioning in children and adolescents. Respondents are asked to indicate the frequency of each symptom on a three-point Likert scale with options of 0 = *never*, 1 = *sometimes*, and 2 = *often*, for a total score ranging from 0 to 70. Multiple studies have documented the reliability and validity of the PSC in samples representing a diversity of age, race, socioeconomic status, and diagnostic conditions [41–43]. These studies show the measure to have strong internal consistency (Cronbach’s alpha = .87–.89; [44, 45]) and test-retest reliability ($r = .61$; [46]). Total PSC scores have been used in a number of studies to assess intervention outcomes [47, 48], monitor response to treatment [49, 50], and track children’s functioning over time [24, 51]. A recent study demonstrated strong agreement between positive PSC screens and functional impairment [52], highlighting the usefulness of the measure in tracking outcomes important to children and their families.

Factor analysis of the PSC items has shown that the measure loads onto three factors that capture specific dimensions of problem behavior in children (internalizing, externalizing, and attention) [28, 53, 54]. Each factor has been operationalized in a subscale that includes five or

seven items and, like the global scale, can be dichotomized into case/not case categories based on a validated cutoff score. Gardner and colleagues [31] reported relatively high rates of agreement between the three PSC subscales and the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children – Present and Lifetime version as well as some of the most commonly used outcome measures in child psychiatry, specifically the CDI and the SCARED.

The Internalizing Subscale of the PSC (PSC-IS) examined in the current study is comprised of the five items identified by Gardner et al. [28] that assess symptoms of depression and anxiety (see Fig. 1). A total score of five or higher indicates impairment in this area.

Clinician-Completed Measures

Intake scores on the clinician-completed BPRS-C and CGAS were used for comparison to PSC scores. The BPRS-C consists of 21 distinct symptom areas, each rated for severity on a seven-point Likert scale based on the clinician's interview. Total scores on the measure range from 0 to 126 and are computed by summing the items, with higher scores indicating higher symptom severity [55]. Changes in BPRS-C total scores have been used to evaluate treatment effectiveness in community [56] and residential [57] treatment settings. The BPRS-C items are nested in seven subscales of three items, each representing specific problem areas [39]. In 2001, Lachar and colleagues published an article validating a second-order Internalization Scale of the BPRS-C (BPRS-C-IS) [55]. A patient's score on the BPRS-C-IS is the sum of his/her scores on the Depression, Anxiety, and Thinking Disturbance Subscales, therefore providing an appropriate

comparison for the PSC-IS. We used the BPRS-C-IS to evaluate the relationship between parent- and clinician-rated internalizing symptoms.

The CGAS is a numeric scale ranging from 1 to 100. It is virtually identical to the Axis V rating in the DSM-IV and is used by clinicians to rate the overall functioning of children, with lower scores indicating poorer functioning [40]. It has been widely used in both research and clinical care [41].

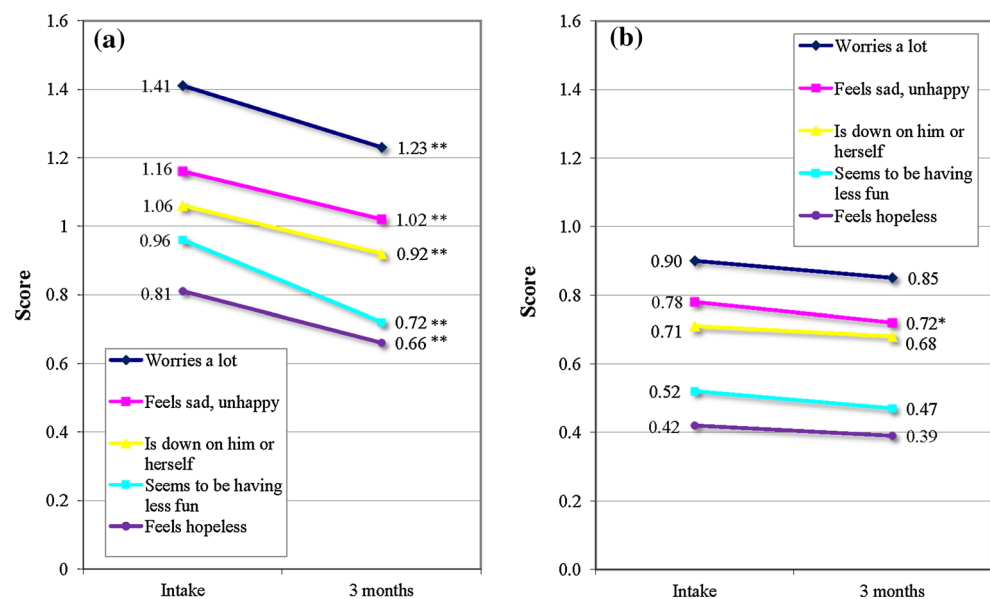
DSM-IV Diagnosis

Two experienced clinicians sorted all DSM-IV diagnoses used into two mutually exclusive categories (internalizing versus non-internalizing). Initial classifications were performed by each clinician separately, after which the two lists were compared for the presence of inconsistencies. Classification was nearly unanimous across both types of diagnoses and, for the small number of disorders in which the classifications did differ ($n = 3$), the two clinicians discussed each case individually until clinical consensus was reached. The list of DSM-IV internalizing diagnoses is presented in the Appendix. A member of the study team (blind to PSC scores) reviewed the diagnoses given to all patients and coded them as either internalizing or non-internalizing on the basis of the clinicians' categorization. Patients were classified as internalizing if their clinician indicated an internalizing disorder as either their primary or secondary diagnosis.

Data Analysis

Initial analyses used Chi square and analysis of variance (ANOVA) to compare internalizing and non-internalizing

Fig. 1 **a** Mean scores on PSC internalizing subscale items at intake and 3-month follow-up for patients with internalizing diagnoses. **b** Mean scores on PSC internalizing subscale items at intake and 3-month follow-up for patients with non-internalizing diagnoses



groups in our first analytic sample on gender, age, insurance type, treatment modality, and standardized measure scores to ascertain whether there were differences between the groups prior to treatment. We then evaluated the level of agreement between clinician diagnoses and parent ratings of patients' internalizing risk at intake.

The remaining analyses were conducted using our second analytic sample of patients with complete data at 3 months into treatment. PSC-IS total and individual item scores from intake and 3-month follow-up were analyzed cross-sectionally and longitudinally for patients with internalizing and non-internalizing diagnoses. Significance of the change from intake to 3 months was ascertained using paired *t*-tests, and differences in the amount of change for internalizing and non-internalizing patients compared using one-way ANOVAs. Effect sizes are presented as a measure of the strength of the observed changes and reported as Cohen's *d* [58].

To further evaluate treatment according to the PSC-IS, we used the RCI to calculate rates of reliable and clinically significant change for patients with internalizing and non-internalizing diagnoses. According to the formula proposed by Jacobson and Truax [33], the RCI for the PSC-IS is two points. Therefore, a change of at least two points on the PSC-IS reflects psychometrically reliable improvement or deterioration in internalizing symptoms. Patients were defined as having experienced clinically significant improvement if they had initial PSC-IS scores above the cutoff (greater than or equal to five), reliable improvement (score decreased by two or more points), and three-month follow-up scores below the cutoff (less than five). This level of improvement suggested a positive and significant response to treatment as well as a level of psychological health similar to that of non-patients. Conversely, patients were defined as showing clinically significant deterioration if they had initial PSC-IS scores below the cutoff (less than five), reliable deterioration (score increased by two or more points), and three-month follow-up scores above the cutoff (greater than or equal to five). This level of deterioration suggested a negative and significant response to treatment that also caused the patient to be newly classified as at-risk. These measures were used to evaluate the number of patients who had achieved remission of their symptoms 3 months into treatment and whether this differed between our two analytic subsamples.

In order to validate the reliable and clinically significant change metrics for the PSC-IS, we compared these categorical changes on the PSC-IS to changes in mean scores on the clinician-rated BPRS-C-IS and CGAS. We also calculated the correlations between change in parent-reported internalizing symptoms (on the PSC-IS) and clinician-reported change in internalizing symptoms (BPRS-C-IS) and overall functioning (CGAS) over 3 months. Lastly, we statistically compared rates of improvement and

deterioration by treatment modality in order to ensure that type of care (e.g., medication versus therapy) did not confound our findings.

Results

Participant Characteristics at Intake

Of the 2,932 eligible youth who entered outpatient treatment and were assessed between August 1, 2007 and July 31, 2013, a clinician-completed outcome form was collected for 2,169 patients (74 %). Among those patients, 1,692 patients (78 %) had a PSC form completed by their parent before their intake appointment. To ensure restricting our sample to patients with parent-completed forms did not significantly bias our results, we statistically compared scores on clinician-rated measures for patients with and without PSCs; these analyses confirmed that the samples did not differ according to scores on the two clinician reports (mean BPRS-C-IS score for patients with PSC = 9.59 vs. no PSC = 9.87; $F = .58$, $p > .05$ and mean CGAS score for patients with PSC = 57.83 vs. no PSC = 57.74; $F = .04$, $p > .05$).

Fifty-three percent (900) of the 1,692 patients with parent and clinician data had a primary and/or secondary DSM-IV Axis I internalizing diagnosis and 41 % (693) were diagnosed with a non-internalizing DSM-IV Axis I disorder. Anxiety disorder not otherwise specified (NOS, 21.2 %, $n = 192$), mood disorder NOS (18.0 %, $n = 169$), and obsessive-compulsive disorder (12.7 %, $n = 114$) were the most common primary diagnoses in the internalizing group, while attention-deficit/hyperactivity disorder (32.7 %, $n = 226$), adjustment disorder NOS (8.5 %, $n = 59$), and pervasive developmental disorder (5.6 %, $n = 39$) were most common in the non-internalizing group. These 1,593 patients constituted our first analytic sample. The remaining 6 % ($n = 99$) of patients were excluded from the sample because they were missing DSM-IV Axis I diagnoses on their intake forms.

Our second analytic sample included patients with complete PSC data at 3 months into treatment. Thirty-nine percent ($n = 620$) of patients in the first analytic sample were included in this sample. As highlighted in an earlier paper [21], about half of patients lost between intake and 3 months are those who do not enter treatment at MGH Child and Adolescent Psychiatry. These patients are often seen for one or two consultation visits and then referred to more conveniently located treatment facilities, as a result, follow-up forms are not completed by parents or clinicians. Among patients who actually entered treatment in the Child and Adolescent Psychiatry Service, we received follow-up PSCs from 69 % of parents. There were no significant differences in demographic information

between patients with outcome forms at intake only and those with forms at intake and follow-up.

As shown in Table 1, in the current study, patients with internalizing diagnoses were more likely than patients with non-internalizing diagnoses to be female ($\chi^2 = 21.08, p < .001$). These patients were also significantly older ($F = 105.99, p < .001$), less likely to use public insurance ($\chi^2 = 3.87, p < .05$), and more likely to have treatment plans that included medication alone or combined medication and therapy ($\chi^2 = 40.83, p < .001$) than their non-internalizing counterparts. According to parent and clinician measures completed at intake, patients with internalizing diagnoses also had more internalizing symptoms (PSC-IS: $F = 220.73, p < .001$, BPRS-C-IS: $F = 199.34, p < .001$) and poorer functioning overall (PSC: $F = 55.91$, BPRS-C: $F = 49.17$, CGAS: $F = 68.50$, all $p < .001$) than patients with non-internalizing diagnoses.

Agreement between Diagnoses and PSC-IS Scores at Intake

Using the established PSC-IS cutoff score of five or higher, 49 % ($n = 785$) of youth in the first analytic sample screened positive for internalizing problems. Overall, clinician-reported diagnoses and PSC-IS case classifications agreed for 64 % of patients ($\chi^2 = 122.52, p < .001$).

Change Scores on the PSC-IS for Patients with Internalizing and Non-internalizing Diagnoses

Scores on the PSC-IS were used to track outcomes over the first 3 months of treatment and compare internalizing symptom change for patients with internalizing and non-internalizing diagnoses. Among patients in our second analytic sample ($n = 620$), those with internalizing disorders (60 % of total) had significantly higher PSC-IS scores

Table 1 Patient demographics at intake

	Internalizing diagnosis	Non-internalizing diagnosis
No. (%) of sample	900 (56.5)	693 (43.5)
Gender***		
No. (%) male	452 (50.2)	428 (61.8)
No. (%) female	448 (49.8)	265 (38.2)
Mean (SD) age***	12.09 (3.81)	10.03 (4.14)
Insurance type*		
No. (%) commercial/other	774 (86.0)	571 (82.4)
No. (%) Medicaid	126 (14.0)	122 (17.6)
Treatment modality ^{a,***}		
No. (%) therapy only	270 (32.1)	277 (45.3)
No. (%) psychopharmacology only	414 (49.2)	280 (45.8)
No. (%) combination treatment	157 (18.7)	54 (8.8)
Mean (SD) PSC scores		
Global Scale***	26.86 (12.07)	22.36 (11.69)
Internalizing Subscale***	5.36 (2.77)	3.35 (2.54)
Mean (SD) BPRS-C scores		
Global Scale ^{b,***}	20.57 (12.09)	16.26 (11.70)
Internalization Scale ^{c,***}	11.64 (6.18)	7.19 (5.95)
Mean (SD) CGAS score ^{d,***}	56.38 (7.61)	59.64 (7.73)

SD standard deviation, PSC Pediatric Symptom Checklist, BPRS-C Brief Psychiatric Rating Scale for Children, CGAS Children's Global Assessment Scale

* $p < .05$; *** $p < .001$

^a Data on treatment type were only available for 1,452 patients, including 841 with internalizing diagnoses and 611 with non-internalizing diagnoses. Percentages are out of 841 and 611 for the internalizing and non-internalizing groups, respectively

^b Data on the BPRS-C Global Scale were available for 1,533 patients

^c Data on the BPRS-C Internalization Scale were available for 1,516 patients

^d Data on the CGAS were available for 1,535 patients

at intake ($F = 95.19, p < .001$) and 3 months ($F = 46.77, p < .001$) than those with non-internalizing disorders. The average change on the PSC-IS between intake and 3 months was $-.83$ points ($SD = 2.43$) for internalizing patients and $-.22$ points ($SD = 1.94$) for non-internalizing patients. This improvement reached statistical significance only for the internalizing group ($t = 6.57, p < .001$). Between-group comparison similarly showed that the amount of change was significantly greater for internalizing patients ($F = 11.03, p < .001$), with an effect size that was moderate for the internalizing group ($d = .34$) but small for the non-internalizing group ($d = .11$).

Mean scores for all five PSC-IS items decreased significantly ($p < .001$) among patients with internalizing diagnoses (Fig. 1a), whereas the mean score on only one item decreased significantly (feels sad, unhappy; $p < .05$) among patients with non-internalizing diagnoses (Fig. 1b). The magnitude of the difference in change scores between the two subsamples was statistically significant ($p < .05$ to $p < .01$) for four of the PSC-IS items. The difference did not reach significance on the fifth item, “feels sad, unhappy” ($p > .05$).

Reliable and Clinically Significant Change on the PSC-IS

The top two rows of Table 2 compare the rates of reliable and clinically significant change on the PSC-IS for patients with internalizing and non-internalizing diagnoses. Among patients with internalizing diagnoses, 36 % ($n = 132$) experienced reliable improvement on the PSC-IS over the first 3 months of treatment. Sixty-four of those patients

(17 % of the internalizing group) had scores that also crossed the risk threshold, indicating clinically significant improvement. About half (51 %, $n = 190$) of internalizing patients had PSC-IS change scores ranging from -1 to $+1$, which are not meaningful according to the RCI. On the other end of the spectrum, a small number (13 %, $n = 48$) of internalizing patients experienced reliable deterioration. For an even smaller minority (5 %, $n = 18$), this deterioration was clinically significant.

As shown in the second row of Table 2, patients with non-internalizing diagnoses were significantly less likely to experience clinically significant change on the PSC-IS over 3 months than patients with internalizing diagnoses ($\chi^2 = 26.57, p < .001$). While small percentages of non-internalizing patients experienced clinically significant changes [9 % ($n = 22$) improved and 6 % ($n = 14$) deteriorated], the majority (70 %, $n = 176$) did not show reliable change in either direction.

Comparisons with Clinician-Reported Changes in Internalizing Symptoms

We assessed the criterion validity of the PSC-IS by comparing mean change scores on the clinician-rated BPRS-C-IS across levels of change on the PSC-IS. For the entire sample ($n = 583$), the correlation between change scores on the PSC-IS ($M_{\text{change}} = -.58, SD = 2.26$) and BPRS-C-IS ($M_{\text{change}} = -1.54, SD = 6.46$) over the first 3 months of treatment was of moderate size ($r = .38, p < .10$).

As shown in the third row of Table 2, patients who experienced clinically significant improvement according

Table 2 Rates of reliable and clinically significant change on the Pediatric Symptom Checklist Internalizing Subscale by diagnostic category and comparison with changes on clinician-completed measures

	N	Level of change on the PSC-IS				
		Clinically significant improvement ^a	Reliable improvement only ^b	No reliable change ^c	Reliable deterioration only ^d	Clinically significant deterioration ^e
Patient diagnosis***						
No. (%) with internalizing diagnosis	370	64 (17.3)	68 (18.4)	190 (51.4)	30 (8.1)	18 (4.9)
No. (%) with non-internalizing diagnosis	250	22 (8.8)	22 (8.8)	176 (70.4)	15 (6.0)	14 (5.6)
Changes on clinician-completed measures						
Mean (SD) change on BPRS-C-IS***	579	-5.14 (6.87)	-3.89 (7.43)	-.96 (4.53)	2.77 (10.57)	2.10 (6.52)
Mean (SD) change on CGAS***	576	5.06 (10.06)	1.99 (10.77)	1.50 (5.69)	-1.18 (8.26)	-3.74 (13.92)

SD standard deviation, PSC-IS Pediatric Symptom Checklist Internalizing Subscale, BPRS-C-IS Brief Psychiatric Rating Scale for Children Internalization Scale, CGAS Children’s Global Assessment Scale

*** $p < .001$

^a Patient’s score decreased by at least 2 points and moved from above to below the cutoff score of 5

^b Patient’s score decreased by at least 2 points but did not move from above to below the cutoff score of 5

^c Patient’s score did not change by more than 1 point

^d Patient’s score increased by at least 2 points but did not move from below to above the cutoff score of 5

^e Patient’s score increased by at least 2 points and moved from below to above the cutoff score of 5

to the PSC-IS also had the most progress on the BPRS-C-IS ($M_{\text{change}} = -5.14$, $SD = 6.87$, $t = 5.72$, $p < .001$, $d = .75$). Those who improved reliably without crossing the risk threshold also improved but to a lesser degree ($M_{\text{change}} = -3.89$, $SD = 7.43$, $t = 4.83$, $p < .001$, $d = .53$). Although post hoc analyses failed to show a significant difference between the reliable and clinically significant improvement groups, the differences between those two and the other three groups were significant (all $p < .001$). Overall, patients who did not experience reliable change on the PSC-IS improved slightly on the BPRS-C-IS ($M_{\text{change}} = -.96$, $SD = 4.53$, $t = 3.89$, $p < .001$, $d = .21$), while those who deteriorated reliably ($M_{\text{change}} = 2.77$, $SD = 10.57$, $t = -1.72$, $p = .09$, $d = .27$) or clinically significantly ($M_{\text{change}} = 2.10$, $SD = 6.52$, $t = -1.76$, $p = .09$, $d = .32$) also had worse scores on the measure 3 months into treatment.

Comparisons with Clinician-Reported Changes in Overall Functioning

A second clinician-rated measure provided additional evidence of the PSC-IS's validity by demonstrating parent and clinician agreement on the patient's global functioning. Among all patients with complete data on the parent-completed PSC-IS and clinician-completed CGAS ($n = 576$), the correlation between change scores on the measures was $r = -.24$ ($p < .01$).

As shown in the bottom row of Table 2, patients who achieved clinically significant improvement on the PSC-IS had the greatest mean increase in CGAS score ($M_{\text{change}} = 5.06$, $SD = 10.06$, $t = 4.48$, $p < .001$, $d = .50$). Those who experienced reliable improvement alone also improved on the CGAS ($M_{\text{change}} = 1.99$, $SD = 10.77$, $t = -1.69$, $p = .09$, $d = .19$), although the group's mean change only trended towards significance and was significantly smaller than that of the clinically significant improvement group ($p < .05$). Consistent with change on the BPRS-C-IS, improvement on the CGAS for patients who did not experience reliable change on the PSC-IS was non-significant ($M_{\text{change}} = 1.50$, $SD = 5.69$, $t = 1.66$, $p = .10$, $d = .21$). Although clinically significant deterioration on the PSC-IS was associated with a greater mean decline on the CGAS than reliable deterioration alone ($M_{\text{change}} = -3.74$, $SD = 13.92$, $t = 1.40$, $p = .17$, $d = .29$ versus $M_{\text{change}} = -1.18$, $SD = 8.26$, $t = .90$, $p = .37$, $d = .14$, respectively), this difference did not reach significance.

Rates of Change by Treatment Modality

In order to ensure that treatment modality did not confound our results, we statistically compared change on the PSC-IS for patients whose treatment plans on their intake forms

included medication alone, therapy alone, or combined medication and therapy. These analyses confirmed that rates of improvement and deterioration were similar across treatment types. Among patients with internalizing diagnoses, between-group comparisons of PSC-IS scores showed that treatment modality did not predict significant differences in rates of clinically significant improvement ($F = .96$, $p = .40$), reliable improvement ($F = .25$, $p = .78$), no reliable change ($F = .13$, $p = .87$), reliable deterioration ($F = .11$, $p = .90$), or clinically significant deterioration ($F = .17$, $p = .85$). The lack of significant differences between treatment modality suggested that rates of improvement and deterioration were of the same magnitude regardless of the type of treatment received.

Discussion

The results of this study suggest that the PSC-IS can be a useful screen and/or treatment outcome measure when anxiety or depressive problems are of concern in child psychiatry or pediatrics. At intake in an outpatient child psychiatry clinic, continuous and categorical cut-off scores on the PSC-IS indicated agreement between parents' reports of internalizing symptoms and clinicians' diagnoses of internalizing disorders, a finding which supported the measure's construct validity in a real-world sample.

Among patients who remained in treatment for 3 months, change scores on the PSC-IS and its individual items were significantly larger among patients with internalizing diagnoses than among those with non-internalizing diagnoses. This higher rate of improvement in the internalizing group demonstrated the responsiveness of the PSC-IS to the indicators it is designed to measure, thereby suggesting it could be used to track symptoms of anxiety and depression in youth over the course of treatment. In this way, the PSC-IS is comparable to widely-used, diagnosis-specific measures like the CDI [15] and the SCARED [16]. Yet, unlike those measures, the PSC-IS is also part of a global measure that assesses other types of problems (e.g., attention, externalizing). Clinicians can therefore utilize the PSC to quantify changes in both overall functioning and multiple, disorder-specific domains. Furthermore, the PSC's orientation towards functioning is familiar to primary care pediatricians. The PSC subscales are already being used routinely in primary care as a screen for referring children with externalizing problems to a psychosocial treatment program [59–62] and by the Chilean government to identify first grade students who may need a referral to mental health care for attention and/or internalizing problems [24, 63].

According to the reliable and clinically significant change metrics, 36 % of patients with internalizing

diagnoses experienced statistically reliable improvement on the PSC-IS over the first 3 months of treatment in a child psychiatry clinic. Nearly half of those patients began treatment in the clinical range but, after 3 months, had scores typical of healthy individuals. According to Jacobson and Truax's conceptualization of clinically significant change, these were patients who had responded positively and significantly to treatment and recovered from their current episode of psychiatric illness [33]. As clinically significant improvement may signify a remission of the patient's internalizing symptoms, this may be an opportune point for clinicians to review or reevaluate their current treatment plan. On the opposite end of the spectrum, 13 % of internalizing patients reliably or clinically significantly deteriorated on the PSC-IS over their first 3 months, a pattern of response that suggests the need to reexamine and possibly redirect the approach taken with such patients. Reasons for treatment decline are varied; symptom increase, breakdown in patient-clinician communication, and a failure to adhere to the treatment protocol are just a few of the reasons why symptoms may worsen during the course of therapy [64]. Since therapists are notably inaccurate when it comes to predicting which of their patients are likely to deteriorate, regular use of simple outcome measures like the PSC could be a significant aid to clinical decision making and treatment planning. Moreover, as patients who experience treatment decline or failure are likely to require a greater amount of treatment resources or show a worse trajectory of overall functioning and well-being [64], identifying these cases early on through the use of a systematic screening process may help prevent these negative outcomes.

Significantly fewer patients with non-internalizing diagnoses experienced reliable or clinically significant change on the PSC-IS, a finding consistent with the assumption that their treatments were focused on other areas. At the same time, the achievement of reliable or clinically significant improvement by some patients with non-internalizing diagnoses suggests that treatment effects and/or the diagnoses given may have been non-specific.

Comparisons between longitudinal scores on the parent-completed PSC-IS and clinician-completed BPRS-C-IS and CGAS provided further validation for the PSC-IS. The moderate correlations between change scores on parent and clinician measures suggested good agreement on patients' internalizing symptoms and overall functioning over 3 months. At the same time, the existence of some disparity between the two assessments highlights the importance of using both parent and clinician measures as these may provide unique information about patient progress. Change scores on the BPRS-C-IS and CGAS also supported the criterion validity of the PSC-IS, as clinically significant improvement and deterioration on the PSC-IS

were both associated with statistically significant mean changes in the predicted directions on the two distinct clinician measures. Similarly, evidence for convergent validity of the PSC-IS was provided by both continuous scores on the measure, which correlated with scores on the clinician-completed internalizing subscale, and its categorical cutoff, which showed a high rate of agreement with clinical diagnosis. On the other hand, small and generally non-significant change among patients without an internalizing diagnosis offered preliminary support for the measure's divergent validity. Future work comparing change on this and the PSC's other symptom-specific subscales will help to further illuminate its psychometric properties.

One noteworthy finding in this study concerned its smaller effect sizes compared to those reported in a number of clinical trials [65, 66]. For example, patients treated with combined medication and therapy for 3 months had a mean treatment effect size of Hedges' $g = .44$ on the BPRS-C-IS [66]. In the Treatment for Adolescent Depression Study (TADS) [65], combined fluoxetine and cognitive-behavioral therapy had an effect size of Hedges' $g = .98$ over the same time period on the clinician-reported Children's Depression Rating Scale-Revised. Although there are many important differences between the current study and TADS (e.g., the sample for TADS included older, treatment naïve youth with non-comorbid major depression), the fact that the current real-world sample can be compared to a milestone clinical trial and obtained a result of a comparable order of magnitude (albeit substantially smaller) further supports the potential utility of having data from brief, standardized measures used routinely in clinical settings. It is also important to point out that since child psychiatry clinics in tertiary care hospitals are not usually the initial point-of-contact when mental health issues arise, new patients at these clinics have often already received treatment in locations such as school or primary care. This may reduce the amount of improvement those patients show when they begin specialty care.

The current study had several limitations, including the fact that it was conducted in a single outpatient child psychiatry clinic at an academic medical center. Future research is tasked with determining the generalizability of these findings to community mental health and pediatric settings. In addition, due to the study's naturalistic design, a majority of patients were missing follow-up data, and we cannot discount the possibility of selection bias that might have been incurred by using only patients with repeat forms. However, our follow-up rate was consistent with other real-world psychiatry samples [67, 68], even with a relatively long (three-month) interval between patients' first and second assessments. As mentioned earlier, a large percentage of this clinic's patients are seen for treatment

consultation specifically and therefore only provide baseline information. Future aims for this outcomes project are to consider using shorter intervals between PSC assessments in order to examine the trajectories of improvement and deterioration over a larger number of time points and for a larger proportion of patients. Studies already in progress will follow the current sample out to 6 and 9 months of treatment. Furthermore, applying the statistical criteria used here to the PSC's externalizing and attention subscales and evaluating their effectiveness in identifying and monitoring those symptoms are additional aims of this research project.

As healthcare increases its use of information systems and places greater emphasis on patient engagement and demonstrating treatment effectiveness, especially within Accountable Care Organizations, the regular administration of questionnaires to assess degree of impairment and to track progress, adherence to treatment protocols, costs, and quality will become the norm rather than the exception. The approach and results of this study are one step towards these ends as well as towards the goal of making the routine psychiatric care of children and adolescents more reliable and effective.

Summary

Previous research has shown that the PSC can identify improvement and deterioration in overall functioning over the first several months of treatment [32], and the current paper suggests that it can be an effective measure of changes within the specific area of internalizing problems as well. The PSC could therefore be a useful measure for referring pediatricians as well as for child psychiatrists and psychologists in that it is brief, offers a preliminary assessment of a patient's functioning globally and in symptom-specific domains, and can effectively track early response to treatment.

Acknowledgments A portion of Dr. Murphy's time for the planning and data analysis for this study was made possible by a grant from the Fuss Family Fund. The support of the Fuss Family Fund is hereby gratefully acknowledged.

Conflict of interest None declared.

Appendix: DSM-IV Axis I Diagnoses Coded as Internalizing Disorders

296.00 Bipolar I Disorder, Single Manic Episode, Unspecified
 296.20 Major Depressive Disorder, Single Episode, Unspecified

296.21 Major Depressive Disorder, Single Episode, Mild
 296.22 Major Depressive Disorder, Single Episode, Moderate
 296.23 Major Depressive Disorder, Single Episode, Severe Without Psychotic Features
 296.30 Major Depressive Disorder, Recurrent, Unspecified
 296.31 Major Depressive Disorder, Recurrent, Mild
 296.32 Major Depressive Disorder, Recurrent, Moderate
 296.33 Major Depressive Disorder, Recurrent, Severe Without Psychotic Features
 296.34 Major Depressive Disorder, Recurrent, Severe With Psychotic Features
 296.35 Major Depressive Disorder, Recurrent, In Partial Remission
 296.40 Bipolar I Disorder, Most Recent Episode Manic, Unspecified
 296.42 Bipolar I Disorder, Most Recent Episode Manic, Moderate
 296.50 Bipolar I Disorder, Most Recent Episode Depressed, Unspecified
 296.53 Bipolar I Disorder, Most Recent Episode Depressed, Severe Without Psychotic Features
 296.60 Bipolar I Disorder, Most Recent Episode Mixed, Unspecified
 296.64 Bipolar I Disorder, Most Recent Episode Mixed, Severe With Psychotic Features
 296.70 Bipolar I Disorder, Most Recent Episode Unspecified
 296.80 Bipolar Disorder NOS
 296.89 Bipolar II Disorder
 296.90 Mood Disorder NOS
 300.00 Anxiety Disorder, NOS
 300.01 Panic Disorder without Agoraphobia
 300.02 Generalized Anxiety Disorder
 300.20 Phobia Unspecified
 300.21 Panic Disorder with Agoraphobia
 300.22 Agoraphobia without History of Panic Disorder
 300.23 Social Phobia
 300.29 Specific Phobia
 300.30 Obsessive Compulsive Disorder
 300.40 Dysthymic Disorder
 308.30 Acute Stress Disorder
 309.81 Posttraumatic Stress Disorder
 311.00 Depressive Disorder NOS

References

1. Williams SB, O'Connor EA, Eder M, Whitlock EP (2009) Screening for child and adolescent depression in primary care settings: a systematic evidence review for the US Preventive Services Task Force. *Pediatrics* 123:e716–e735
2. Kessler RC, Walters EE (1998) Epidemiology of DSM-III-R major depression and minor depression among adolescents and

- young adults in the National Comorbidity Survey. *Depress Anxiety* 7:3–14
3. Woodward LJ, Fergusson DM (2001) Life course outcomes of young people with anxiety disorders in adolescence. *J Am Acad Child Adolesc Psychiatry* 40:1086–1093
 4. Murray C, Lopez A (1996) *The global burden of disease*. Harvard University Press, Cambridge
 5. Fergusson DM, Woodward LJ (2002) Mental health, educational, and social role outcomes of adolescents with depression. *Arch Gen Psychiatry* 59:225–231
 6. Pine DS, Cohen P, Gurley D, Brook J, Ma Y (1998) The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Arch Gen Psychiatry* 55:56–64
 7. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT (2008) Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* 5:e45
 8. Duncan LE, Keller MC (2011) A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am J Psychiatry* 168:1041–1049
 9. Murphy JM, McCarthy AE, Baer L, Zima BT, Jellinek MS (2014) Alternative national guidelines for treating attention and depression problems in children: comparison of treatment approaches and prescribing rates in the United Kingdom and United States. *Harv Rev Psychiatry*. doi:10.1097/HRP.000000000000026 (Published online April 14, 2014)
 10. Batty MJ, Moldavsky M, Foroushani PS, Pass S, Marriott M, Sayal K, Hollis C (2013) Implementing routine outcome measures in child and adolescent mental health services: from present to future practice. *Child Adolesc Ment Health* 18:82–87
 11. Bickman L, Kelley SD, Breda C, de Andrade AR, Riemer M (2011) Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. *Psychiatr Serv* 62:1423–1429
 12. Hall CL, Moldavsky M, Taylor J, Sayal K, Marriott M, Batty MJ et al (2014) Implementation of routine outcome measurement in child and adolescent mental health services in the United Kingdom: a critical perspective. *Euro Child Adolesc Psychiatry* 23:239
 13. Knap C, Koesters M, Schoefer D, Becker T, Puschner B (2009) Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. *Br J Psychol* 195:15–22
 14. Marshall S, Haywood K, Fitzpatrick R (2006) Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract* 12:559–568
 15. Saylor CF, Finch AJ, Spirito A, Bennett B (1984) The children's depression inventory: a systematic evaluation of psychometric properties. *J Consult Clin Psychol* 52:955–967
 16. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J et al (1997) The Screen for Child Anxiety Related Emotional Disorders (SCARED): scale construction and psychometric characteristics. *J Am Acad Child Adolesc Psychiatry* 36:545–553
 17. Kroenke K, Spitzer RL, Williams JB (2001) The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 16:606–613
 18. Kroenke K, Spitzer RL, Williams JBW (2003) The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care* 41:1284–1292
 19. *Addressing Mental Health Concerns in Primary Care: A Clinician's Toolkit* [computer program] (2010) Elk Grove Village, IL: American Academy of Pediatrics
 20. Jellinek MS, Murphy JM, Burns BJ (1986) Brief psychosocial screening in outpatient pediatric practice. *J Pediatr* 109:371–378
 21. Murphy JM, Masek BJ, Babcock R, Jellinek M, Gold J, Drubner S et al (2011) Measuring outcomes in outpatient child psychiatry: the contribution of electronic technologies and parent report. *Clin Child Psychol Psychiatry* 16:146–160
 22. Murphy JM, Kamin HS, Masek B, Vogeli C, Caggiano R, Sklar K et al (2012) Using brief clinician and parent measures to track outcomes in outpatient child psychiatry: longer term follow-up and comparative effectiveness. *Child Adolesc Ment Health* 17:222–230
 23. Hacker KA, Penfold R, Arsenault L, Zhang F, Murphy M, Wisow L (2014) Screening for behavioral health issues in children enrolled in Massachusetts Medicaid. *Pediatrics* 133:46–54
 24. Murphy JM, Guzmán J, McCarthy AE, Squicciarini AM, George M, Canenguez K (2014) Mental health predicts better academic outcomes: a longitudinal study of elementary school students in Chile. *Child Psychiatry Hum Dev*. doi:10.1007/s10578-014-0464-4 (Published online April 26, 2014)
 25. Zima BT, Murphy JM, Scholle SH, Hoagwood KE, Sachdeva RC, Mangione-Smith R, Woods D, Kamin HS, Jellinek MS (2013) National quality measures for child mental health care: background, progress, and next steps. *Pediatrics* 131:S38–S49
 26. Conway PH, Mostashari F, Clancy C (2013) The future of quality measurement for improvement and accountability. *JAMA* 309:2215–2216
 27. Delgadillo J, McMillan D, Leach C, Lucock M, Gilbody S, Wood N (2014) Benchmarking routine psychological services: a discussion of challenges and methods. *Behav Cogn Psychother* 42:16–30
 28. Gardner W, Murphy JM, Childs G, Kelleher K, Pagano M, Jellinek M (1999) The PSC-17: a brief Pediatric Symptom Checklist with psychosocial problem subscales. A report from PROS and ASPN. *Ambul Child Health* 5:225–236
 29. American Psychiatric Association (2000) *Diagnostic and statistical manual of mental disorders*. American Psychiatric Publishing, Arlington Fourth Edition—Text Revision (DSM-IV-TR)
 30. Wren FJ, Bridge JA, Birmaher B (2004) Screening for childhood anxiety symptoms in primary care: integrating child and parent reports. *J Am Acad Child Adolesc Psychiatry* 43:1364–1371
 31. Gardner W, Lucas A, Kolko DJ, Campo JV (2007) Comparison of the PSC-17 and alternative mental health screens in an at-risk primary care sample. *J Am Acad Child Adolesc Psychiatry* 46:611–618
 32. Murphy JM, Blais MA, Baer L, McCarthy A, Kamin H, Masek B (2013) Measuring outcomes in outpatient child psychiatry: reliable improvement, deterioration, and clinically significant improvement. *Clin Child Psychol Psychiatry*. doi:10.1177/1359104513494872 (Published online July 9, 2013)
 33. Jacobson NS, Truax P (1991) Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 59:12–19
 34. Evans C, Margison F, Barkham M (1998) The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evid Based Ment Health* 1:70–72
 35. Lambert MJ, Hansen NB, Finch AE (2001) Patient-focused research: using patient outcome data to enhance treatment effects. *J Consult Clin Psychol* 69:159–172
 36. McMillan D, Gilbody S, Richards D (2010) Defining successful treatment outcome in depression using the PHQ-9: a comparison of methods. *J Affect Disord* 127:122–129
 37. Wood A, Harrington R, Moore A (1996) Controlled trial of a brief cognitive-behavioural intervention in adolescent patients with depressive disorders. *J Child Psychol Psychiatry* 37:737–746
 38. Gallagher HM, Rabian BA, McCloskey MS (2004) A brief group cognitive-behavioral intervention for social phobia in childhood. *J Anxiety Disord* 18:459–479
 39. Overall JE, Pfefferbaum B (1982) The Brief Psychiatric Rating Scale for Children. *Psychopharmacol Bull* 18:10–16

40. Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H et al (1983) A children's global assessment scale. *Arch Gen Psychiatry* 40:1228–1231
41. Gold J, Buonopane R, Caggiano RA, Picciotto M, Vogeli C, Kanner NT et al (2009) Assessing outcomes in child psychiatry. *Am J Manag Care* 15:210–216
42. Bala G, Golubović Š, Milutinović D, Katić R (2012) Psychometric properties of the Pediatric Symptom Checklist in pre-school children in Serbia. *Med Glas* 9:356–362
43. Gardner W, Lucas A, Kolko DJ, Campo JV (2007) Comparison of the PSC-17 and alternative mental health screens in an at-risk primary care sample. *J Am Acad Child Adolesc Psychiatry* 46:611–618
44. Reijneveld SA, Vogels AG, Hoekstra F, Crone MR (2006) Use of the Pediatric Symptom Checklist for the detection of psychosocial problems in preventive child healthcare. *BMC Public Health* 6:197
45. Thun-Hohenstein L, Herzog S (2008) The predictive value of the pediatric symptom checklist in 5-year-old Austrian children. *Eur J Pediatr* 167:323–329
46. Pagano ME, Cassidy LJ, Little M, Murphy JM, Jellinek MS (2000) Identifying psychosocial dysfunction in school-age children: the Pediatric Symptom Checklist as a self-report measure. *Psychol Sch* 37:91
47. Kleinman RE, Hall S, Green H, Patton K, Pagano ME, Murphy JM (2002) Diet, breakfast, and academic performance in children. *Ann Nutr Metab* 46(Suppl 1):24–30
48. Stein BD, Jaycox LH, Kataoka SH, Wong M, Tu W, Elliott MN (2003) A mental health intervention for schoolchildren exposed to violence: a randomized controlled trial. *JAMA* 290:603–611
49. Furudate N, Komada Y, Kobayashi M, Nakajima S, Inoue Y (2014) Daytime dysfunction in children with restless legs syndrome. *J Neurol Sci* 336:232–236
50. Terrone G, Parente I, Romano A, Auricchio R, Greco L, Del Giudice E (2013) The Pediatric Symptom Checklist as screening tool for neurological and psychosocial problems in a paediatric cohort of patients with coeliac disease. *Acta Paediatr* 102:e325–e328
51. Hacker KA, Williams S, Mysgmarjav S, Cabral H, Murphy M (2009) Persistence and change in Pediatric Symptom Checklist scores over 10–18 months. *Acad Pediatr* 9:270–277
52. Blucker RT, Jackson D, Gillaspay JA, Hale J, Wolraich M, Gillaspay SR (2014) Pediatric behavioral health screening in primary care: a preliminary analysis of the Pediatric Symptom Checklist-17 with functional impairment items. *Clin Pediatr* 53:449–455
53. Reed-Knight B, Hayutin LG, Lewis JD, Blount RL (2011) Factor structure of the pediatric symptom checklist with a pediatric gastroenterology sample. *J Clin Psychol Med Settings* 18: 299–306
54. Stoppelbein L, Greening L, Moll G, Jordan S, Suozzi A (2012) Factor analyses of the Pediatric Symptom Checklist-17 with African-American and Caucasian pediatric populations. *J Pediatr Psychol* 37:348–357
55. Lachar D, Randle SL, Harper RA, Scott-Gurnell KC, Lewis KR, Santos CW (2001) The Brief Psychiatric Rating Scale for Children (BPRS-C): validity and reliability of an anchored version. *J Am Acad Child Adolesc Psychiatry* 40:333–340
56. Plante TG, Couchman CE, Diaz AR (1995) Measuring treatment outcome and client satisfaction among children and families. *J Ment Health Adm* 22:261–269
57. Henderson L, McIlhane K, Wasser T (2010) Measuring outcomes of multiple diagnosis groups in residential treatment using the Brief Psychiatric Rating Scale for Children (BPRS-C). *Child Youth Serv Rev* 32:98–102
58. Hannan C, Lambert MJ, Harmon C, Nielsen SL, Smart DW, Shimokawa K, Sutton SW (2005) A lab test and algorithms for identifying clients at risk for treatment failure. *J Clin Psychol* 61:155–163
59. Lindhiem O, Kolko DJ (2011) Trajectories of symptom reduction during treatment for behavior problems in pediatric primary-care settings. *Adm Policy Ment Health* 38:486–494
60. Kolko DJ, Campo JV, Kelleher K, Cheng Y (2010) Improving access to care and clinical outcome for pediatric behavioral problems: a randomized trial of a nurse-administered intervention in primary care. *J Dev Behav Pediatr* 31:393–404
61. Kolko DJ, Campo JV, Kilbourne AM, Kelleher K (2012) Doctor-office collaborative care for pediatric behavioral problems. *Arch Pediatr Adolesc Med* 166:224–228
62. Kolko DJ, Campo J, Kisbourne AM, Hart J, Sakolsky D, Wisniewski S (2014) Collaborative care outcomes for pediatric behavioral health problems: a cluster randomized trial. *Pediatrics* 133:e981–e992
63. Guzman MP, Jellinek M, George M, Hartley M, Squicciarini AM, Canenguez KM, Kuhlthau KA et al (2011) Mental health matters in elementary school: first-grade screening predicts fourth grade achievement test scores. *Eur Child Adolesc Psychiatry* 20:401–411
64. van Sonsbeek MA, Hutschemaekers GG, Veerman JW, Tiemens BB (2014) Effective components of feedback from Routine Outcome Monitoring (ROM) in youth mental health care: study protocol of a three-arm parallel-group randomized controlled trial. *BMC Psychiatry* 14:3
65. Silva MJ, Wells K, Fairbank J, Burns B, Domino M, McNulty S (2004) Treatment for Adolescents With Depression Study (TADS) Team. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: treatment for Adolescents With Depression Study (TADS) randomized controlled trial. *JAMA* 292:807–820
66. Kendall PC (1994) Treating anxiety disorders in children: results of a randomized clinical trial. *J Consult Clin Psychol* 62:100–110
67. Castonguay LG, Boswell JF, Zack SE, Baker S, Boutselis MA, Chiswick NP et al (2010) Helpful and hindering events in psychotherapy: a practice research network study. *Psychotherapy (Chic)* 47:327–344
68. Blais MA, Sinclair SJ, Baity MR, Worth J, Weiss AP, Ball LA (2012) Measuring outcomes in adult outpatient psychiatry. *Clin Psychol Psychother* 19:203–213