Validation of the Children’s Interview for Psychiatric Syndromes (ChIPS) With Psychiatrically Hospitalized Adolescents

LANCE P. SWENSON, PH.D., CHRISTIANNE ESPOSITO-SMYTHERS, PH.D., JEFFREY I. HUNT, M.D., BETH L.G. HOLLANDER, B.A., JENNIFER DYL, PH.D., CHRISTIE J. RIZZO, PH.D., DOUGLAS L. STEINLEY, PH.D., AND ANTHONY SPIRITO, PH.D.

ABSTRACT

Objective: To examine the concurrent validity of the Children’s Interview for Psychiatric Syndromes (ChIPS) for adolescent inpatients. Method: Participants included 97 adolescents ages 12 to 18 admitted to an adolescent inpatient unit. Participants were administered the ChIPS and the Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime version (present questions only). Participants also completed self-report measures of adjustment (e.g., the Reynolds Adolescent Depression Scale–2). Results: More diagnoses were made with the ChIPS (mean 4.44) compared to the Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime version (mean 3.04; \( p < .001 \)). The percentage of agreement ranged from 59% to 98%. Kappa coefficients indicated agreement ranging from slight for oppositional defiant disorder \((\kappa = .18)\) to substantial for substance use \((\kappa = .66)\); the majority of \(\kappa\) values ranged from .26 to .60. When ChIPS endorsements were examined relative to construct-specific self-report measures of impairment, adolescents diagnosed by the ChIPS with a disorder scored significantly higher than adolescents who were not diagnosed with a disorder. Conclusions: The findings indicate moderate agreement between ChIPS diagnoses and Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime version diagnoses. ChIPS diagnoses appear consistent with self-report measures of adjustment. J. Am. Acad. Child Adolesc. Psychiatry, 2007;46(11):1482–1490. Key Words: Children’s Interview for Psychiatric Syndromes, Schedule for Affective Disorders and Schizophrenia.
because administration and scoring do not require high levels of clinical training (Weller et al., 2000); however, the ChIPS administration manual recommends that the results should always be interpreted by a licensed clinician (Weller et al., 1999).

The ChIPS demonstrates adequate sensitivity, specificity, and concurrent validity in relation to the Diagnostic Interview for Children and Adolescents (DICA-R-C; Welner et al., 1987) and to expert consensus (Weller et al., 2000). It is important to note that comparable results have been obtained using community-based (Fristad et al., 1998b) and clinical (Fristad et al., 1998a; Teare et al., 1998a,b) samples. Several limitations of this research are noteworthy. First, the ChIPS developers have conducted the only research of the interview’s validity; independent investigations of the ChIPS are lacking (for a related discussion, see Weller et al., 2000).

A second limitation of the prior research is the criterion used to evaluate the DSM-IV version of the ChIPS. The ChIPS was originally developed using DSM-III (American Psychiatric Association, 1980) criteria (Teare et al., 1998a); updates were systematically implemented as revisions were made to DSM criteria (Weller et al., 2000). Of the previously mentioned validation studies, only two examined the DSM-IV version of the ChIPS (Fristad et al., 1998a,b). In these studies the sensitivity and specificity of the ChIPS were assessed relative to the DICA-R-C, which uses DSM-III-R criteria (American Psychiatric Association, 1988). Fristad et al. (1998b) state that observed discrepancies “were not obviously attributable” to the different diagnostic criteria of the interviews. Of concern, however, is that the reliability of the DSM-IV version of the ChIPS has not been established with comparable, DSM-IV–based interviews. In addition, multimethod approaches to examining the validity of the interview (e.g., comparing ChIPS diagnoses to self-report measures of symptomatology) are lacking.

Evidence supporting the use of ChIPS with adolescents, particularly adolescent psychiatric populations, was obtained from a relatively small number of participants (i.e., 20 adolescent psychiatric inpatients [Fristad et al., 1998a] and 20 community-based adolescents [Fristad et al., 1998b]). Concerns regarding the limited number of adolescents investigated are somewhat attenuated by the greater number of significant ChIPS and DICA-R-C κ coefficients found among adolescents compared to children (Fristad et al., 1998a). Additional research specifically targeting adolescents, and including significantly larger samples, is needed.

The present study was conducted to address these limitations. Using a sample of adolescent psychiatric inpatients, concurrent validity was assessed with the present (i.e., current) questions of the Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (K-SADS-PL; Kaufman et al., 1997). The K-SADS-PL is ideally suited for this purpose because of the extensive evidence demonstrating the interview’s reliability and validity (Ambrosini, 2000). The validity of the ChIPS was further examined by comparing syndrome endorsement on the ChIPS with construct-specific self-report measures of impairment.

METHOD

Participants

Participants were 97 adolescents (73% female) ages 13 to 18 years (mean 14.97, SD 1.30) admitted to an adolescent inpatient unit of a regional children’s psychiatric hospital in southern New England who also participated in research examining cognitive risk factors for suicidality. This hospital serves all patients regardless of insurance status. According to state census tract data, socioeconomic status (SES) for the population served at this facility includes 16% high SES, 39% middle SES, 15% low SES, and 12% poverty conditions. Self-reported ethnicity of the sample was 80% European American, 5% Native American/American Indian, 4% Latin American, 2% Asian American, 1% African American, and 7% mixed ethnicity/other.

Measures

ChIPS. The ChIPS (Weller et al., 2000) is a highly structured clinical interview that screens for current (i.e., past 2 weeks) symptoms, symptom duration, and impairment for 20 DSM-IV Axis I disorders. Prior validation studies have demonstrated adequate sensitivity and specificity in relation to clinician diagnoses as well as concurrent validity relative to the DICA-R-C (see Weller et al., 2000). The ChIPS child version was used in the present study. Interviews were conducted and scored by one of four highly trained bachelor’s degree–level assistants. Training included viewing and scoring training videotapes, live observation and scoring of interviews, and supervised administration before independent administrations of the interview.

K-SADS-PL (Present Questions Only). The K-SADS-PL (present questions only; Kaufman et al., 1997) is a widely used semistructured diagnostic interview used to assess DSM-IV psychopathology in children and adolescents (Ambrosini, 2000). Prior research with the K-SADS-PL demonstrates excellent test-retest reliability and interrater reliability for major depressive disorder (MDD), generalized anxiety disorder (GAD), conduct disorder, and oppositional defiant disorder (ODD; 0.77–1.00) and adequate reliability for other diagnoses (0.63–0.67; see Kaufman et al., 1997). The child
interview was used in this study, with administration limited to the substance use, mood, disruptive behavior, anxiety disorder, and eating disorders sections. Only current diagnoses (determined by symptom presence, duration, and impairment) were assessed. Interviews were conducted and scored by one of six trained master’s degree- or postdoctoral degree-level clinical psychology trainees. Training included scoring training audiotapes, live observation and scoring of interviews, and supervised administration before independent K-SADS-PL administrations.

Self-Report Measures of Adjustment. Several well-validated, widely used self-report rating scales of psychopathology were administered. These include the Reynolds Adolescent Depression Scale-2 (RADS-2; Reynolds, 2002), a 30-item, 4-point rating scale measuring current depressive symptomatology for adolescents; the Beck Scale for Suicidal Ideation (BSS; Beck and Steer, 1991), a 19-item, 3-point rating scale measuring past-week suicidal ideation; the Suicidal Ideation Questionnaire (SIQ; Reynolds, 1991), a 25-item, 6-point rating scale measuring past-month preoccupation with thoughts of suicide; the Hopelessness Scale for Children (HSC; Kazdin et al., 1986), which includes 17 true-false items measuring feelings of hopelessness and pessimism about the future; the Multidimensional Anxiety Scale for Children-Short Version (MASC-10; March et al., 1997), a 10-item, 4-point rating scale that screens for physical symptoms of anxiety, harm avoidance, social anxiety, and separation/panic; the Screen for Child Anxiety Related Emotional Disorders (SCARED; Birmaher et al., 1997), a 38-item, 3-point rating scale assessing symptoms associated with child anxiety disorder diagnoses; the Trauma Symptom Checklist for Children (TSCC; Briere, 1996), a 55-item, 4-point rating scale measuring trauma symptoms among adolescents; the Aggression Questionnaire (AQ; Buss and Warren, 2000), a 34-item, 5-point rating scale assessing perceived aggression and anger; and the State-Trait Anger Expression Inventory (STAXI; Spielberger, 1988), a 44-item, 4-point rating scale measuring control of and expression of anger.

Procedures
A CONSORT chart detailing participant recruitment and retention is included in Figure 1. Participants were drawn from 467 consecutive admissions (63.8% female; 70.8% European American) to the inpatient facility. As part of the standard admission evaluation, adolescents were interviewed with the ChIPS and were administered the HSC, MASC-10, TSCC, and the STAXI. Results were used for treatment planning. Use of these data was approved for research purposes by the hospital institutional review board.

Of the 467 consecutive admissions, 189 adolescents were invited to participate in research examining cognitive influences on psychopathology and suicidality (demographic information on adolescents who declined participation was not recorded). This...
research was approved by both the university and hospital institutional review boards. Inpatients were eligible for participation if they were 13 to 18 years old, lived with their parents/primary caregivers (who legally had custody) for the past 3 months, were not actively psychotic, and spoke English. Inpatients with a developmental disability or an IQ <70 were excluded. Parental consent and assent were obtained for 138 adolescents (73%). These participants were interviewed with the K-SADS-PL and were administered the RADS-2, BSS, SIQ, SCARED, and AQ.

Of these, 38 participants were not interviewed with the ChIPS due to unexpected discharge or refusal to cooperate with the assessment. In addition, three adolescents administered both interviews were excluded as the length of time between interviews exceeded 60 days. Therefore, 97 participants (21% of the consecutive admissions) who completed both the ChIPS and the K-SADS-PL were retained for the current analyses. Girls were overrepresented in the final sample (i.e., 23.8% of potential females were retained versus 18.2% of potential males [$\chi^2 = 4.67; p < .05$]). Also, 23.8% of the European American adolescents, compared to 9.8% of the minority adolescents, were enrolled in the present study ($\chi^2 = 11.47, p < .001$). No age differences were detected.

The ChIPS and the K-SADS-PL were administered on separate occasions by different interviewers. The ChIPS (part of the intake battery) was always administered before the K-SADS-PL. The interviewers for the K-SADS-PL were blind to the results of the ChIPS and, as research staff, did not have access to any additional information that may have been obtained by unit staff between the administration of the ChIPS and the K-SADS-PL. On average, the two interviews were administered within 3 days of each other (SD 3.60, range 0–16 days).

Data Analysis

Two approaches examined the concurrent validity of the ChIPS. First, standard $\kappa$ coefficients were calculated to assess agreement between the ChIPS and the K-SADS-PL for eight disorders (MDD, GAD, social phobia, PTSD, ADHD, ODD, conduct disorder, and substance use disorder). Low base-rate $\kappa$ coefficients (Verducci et al., 1988) were calculated for four additional disorders: dysthymia, acute stress disorder, anorexia, and bulimia. Both interviews also screened for mania. Bipolar disorder was not diagnosed with the ChIPS at the time of the interview and instead was made during treatment-team case consensus. Because this process relied on several sources of information in addition to the ChIPS, bipolar disorder was not examined.

$\kappa$ accounts for chance-level agreement when estimating interrater reliability (Cohen, 1960); however, $\kappa$ has been criticized as overly conservative (e.g., Petreault and Leigh, 1989). Therefore, the percentage of agreement between the ChIPS and the K-SADS-PL also was examined. Kappa can also be adversely affected by observed prevalence rates as well as by rater/instrument bias (Cicchetti and Feinstein, 1990). To address these concerns, prevalence indices and bias indices were calculated. The prevalence index is the absolute difference between the proportion of cases in which both observers agree that a diagnosis is present and the proportion of cases in which both observers agree that a diagnosis is absent. As the prevalence index approaches unity, chance agreement increases and standard $\kappa$ is reduced. Low base-rate $\kappa$, however, specifically accounts for low prevalence rates (i.e., high proportion of agreement that a diagnosis is absent) when calculating agreement (Verducci et al., 1988).

The bias index is reflected in interrater disagreements on the presence/absence of a diagnosis (i.e., discrepancies in the marginal proportions). Thus, the bias index indicates whether one rater/instrument is more likely to identify a disorder relative to the other. Higher bias indices indicate potentially inflated $\kappa$ values (for further information on the prevalence and bias indices and their effects on $\kappa$ values, see Hoehler, 2000; Sims and Wright, 2005).

The second approach for examining concurrent validity involved one-way analysis of variance testing for mean-level differences in self-reported measures of impairment by ChIPS diagnostic endorsements (i.e., present or absent). Similar tests were conducted by K-SADS-PL diagnostic endorsements for comparison.

RESULTS

Number of Diagnoses

More diagnoses were made with the ChIPS (mean 4.44, SD 2.91) compared to the K-SADS-PL (mean 3.04, SD 1.96), paired $t_{65} = 5.52; p < .001$). The number of ChIPS diagnoses did not differ by sex ($t_{65} = 1.63; p = .11$), age (i.e., older versus younger adolescents; $t_{55} = 1.30; p = .20$), or ethnicity (i.e., white versus nonwhite; $t_{55} = .51; p = .61$). Similarly, the number of K-SADS-PL diagnoses did not differ by sex, age, or ethnicity (all $t \leq 1.67$, all $p \geq .10$).

Concordance Between ChIPS and K-SADS-PL

Prevalence of ChIPS–K-SADS-PL agreements and discrepancies, percentage of agreement, $\kappa$, and prevalence and bias indices are presented in Table 1. Kappa values between .01 and .20 were considered slight agreement, between .21 and .40 as fair agreement, between .41 and .60 as moderate agreement, and between .61 and .80 as substantial agreement (Landis and Koch, 1977). Prevalence indices $\geq .50$ were considered elevated (Hoehler, 2000). Bias indices $\geq .10$ (i.e., reflecting $\geq 10\%$ of the sample diagnosed by one instrument but not the other) were considered elevated. Slight agreement was observed for ODD ($\kappa = .18$; Table 1). This could have been influenced by the ChIPS scoring format. Specifically, differential diagnosis is not considered at the time of scoring. Instead, each disorder is scored in isolation (Weller et al., 1999), without consideration of potential comorbid syndromes. Thus, ODD could have been endorsed by the ChIPS when symptoms occurred exclusively during a depressive episode. Similarly, the ChIPS scoring format allows participants to meet criteria for comorbid ODD and conduct disorder; in both of these circumstances, a diagnosis of ODD would not be appropriate (American Psychiatric Association, 1994). The K-SADS-PL instructs the interviewer to consider differential diagnosis when scoring the results.
<table>
<thead>
<tr>
<th>Syndrome</th>
<th>ChIPS (+)/</th>
<th>ChIPS (+)/</th>
<th>ChIPS (-)/</th>
<th>ChIPS (-)/</th>
<th>% Agreement</th>
<th>κ</th>
<th>Prevalence Index</th>
<th>Bias Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDD</td>
<td>49</td>
<td>21</td>
<td>3</td>
<td>24</td>
<td>75</td>
<td>.49</td>
<td>0.26</td>
<td>0.19</td>
</tr>
<tr>
<td>Dysthymia</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>89</td>
<td>92</td>
<td>.48</td>
<td>0.92</td>
<td>0.02</td>
</tr>
<tr>
<td>GAD</td>
<td>11</td>
<td>18</td>
<td>8</td>
<td>60</td>
<td>73</td>
<td>.29</td>
<td>0.51</td>
<td>0.10</td>
</tr>
<tr>
<td>Social phobia</td>
<td>14</td>
<td>12</td>
<td>11</td>
<td>60</td>
<td>76</td>
<td>.39</td>
<td>0.47</td>
<td>0.01</td>
</tr>
<tr>
<td>PTSD</td>
<td>11</td>
<td>7</td>
<td>12</td>
<td>67</td>
<td>80</td>
<td>.42</td>
<td>0.58</td>
<td>0.05</td>
</tr>
<tr>
<td>Acute stress disorder</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>93</td>
<td>96</td>
<td>.49</td>
<td>0.96</td>
<td>0.02</td>
</tr>
<tr>
<td>Anorexia</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>95</td>
<td>98</td>
<td>.49</td>
<td>0.98</td>
<td>0.02</td>
</tr>
<tr>
<td>Bulimia</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>89</td>
<td>94</td>
<td>.60</td>
<td>0.90</td>
<td>0.06</td>
</tr>
<tr>
<td>ADHD</td>
<td>23</td>
<td>26</td>
<td>10</td>
<td>38</td>
<td>63</td>
<td>.26</td>
<td>0.15</td>
<td>0.16</td>
</tr>
<tr>
<td>ODD</td>
<td>12</td>
<td>37</td>
<td>3</td>
<td>45</td>
<td>59</td>
<td>.18</td>
<td>0.34</td>
<td>0.35</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>20</td>
<td>24</td>
<td>2</td>
<td>51</td>
<td>73</td>
<td>.44</td>
<td>0.32</td>
<td>0.23</td>
</tr>
<tr>
<td>Substance use disorder</td>
<td>19</td>
<td>2</td>
<td>11</td>
<td>65</td>
<td>87</td>
<td>.66</td>
<td>0.47</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Note: + = present/endorsed by the interview; − = absent/not endorsed by the interview; ChIPS = Children’s Interview for Psychiatric Syndromes; K-SADS-PL = Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (present questions only); MDD = major depressive disorder; GAD = generalized anxiety disorder; PTSD = posttraumatic stress disorder; ADHD = attention-deficit/hyperactivity disorder; ODD = oppositional defiant disorder.

Fair agreement was observed for ADHD, GAD, and social phobia (κ = .26–.39; Table 1); however, the elevated bias indices associated with ADHD and GAD (0.16 and 0.10, respectively) reflect the fact that considerably more diagnoses were endorsed by the ChIPS (49 diagnoses of ADHD, 29 diagnoses of GAD) relative to the K-SADS-PL (33 diagnoses of ADHD, 19 diagnoses of GAD).

Moderate agreement was observed for MDD, dysthymia, PTSD, acute stress disorder, anorexia, and bulimia (κ = .42–.60). With regard to PTSD, κ may have been reduced due to the elevated likelihood of not being diagnosed with PTSD (i.e., 70% did not receive a diagnosis of PTSD; prevalence index = 0.58). With regard to conduct disorder and MDD, κ appears unaffected by prevalence; however, the elevated bias indices (0.23 for conduct disorder, 0.19 for MDD) again reflect a tendency for the ChIPS to endorse a greater number of diagnoses (44 diagnoses of conduct disorder, 70 diagnoses of MDD) relative to the K-SADS-PL (22 diagnoses of conduct disorder, 52 diagnoses of MDD).

Substantial agreement was observed for substance use disorder. Although more diagnoses were endorsed by the K-SADS-PL (n = 30) than by the ChIPS (n = 21), the bias index was not significantly elevated.

Comparisons to Self-Reported Adjustment

To further examine the concurrent validity of the ChIPS, one-way analyses of variance were performed to test for mean-level differences in construct-specific self-reported adjustment indices by ChIPS diagnostic endorsements (i.e., present or absent). The self-report questionnaires included measures of depression-related disturbance (RADS-2, BSS, SIQ, and HSC), anxiety symptoms (MASC-10 and SCARED), trauma symptoms (TSCC), and correlates of externalizing behaviors (AQ and STAXI). For comparison purposes, parallel analyses examined K-SADS-PL diagnostic endorsements.
Before conducting these analyses, four diagnostic composites were created by aggregating endorsements of related syndromes. These composites were created because most measures of adjustment assess for general symptoms in these areas versus disorder-specific symptoms and to increase power for analyses. Specifically, diagnoses of MDD and dysthymia were aggregated into a depressive disorder composite (i.e., participants who were diagnosed with either MDD or dysthymia were given a score of 1; participants were given a score of 0 if they were not diagnosed with either disorder). Similarly, diagnoses of GAD and social phobia were aggregated into an anxiety disorder composite, diagnoses of PTSD and acute stress disorder were aggregated into a trauma composite, and diagnoses of ADHD, ODD, and conduct disorder were aggregated into an externalizing composite. Composites were constructed separately for the ChIPS and for the K-SADS-PL.

As evidenced in Table 2, adolescents diagnosed with a depressive disorder by the ChIPS scored significantly higher on the RADS-2, BSS, SIQ, and HSC relative to adolescents not diagnosed with a depressive disorder by the ChIPS. Similarly, significant differences were evident between adolescents with a ChIPS anxiety diagnosis compared with those without an anxiety diagnosis on the MASC-10 and SCARED, between adolescents with a ChIPS trauma-related diagnosis compared with those without a trauma-related diagnosis on the TSCC, and between adolescents diagnosed with a ChIPS externalizing disorder compared with those without an externalizing disorder on the AQ and STAXI. These results, and the magnitude of differences between those diagnosed with a disorder compared to those without, were almost identical to the results obtained with the K-SADS-PL for all nine self-report questionnaires (Table 2).

DISCUSSION

The present study is the first independent examination of the ChIPS and the first to compare the DSM-IV version of the ChIPS to a well-validated interview based on DSM-IV criteria (i.e., the K-SADS-PL). The present study also extends prior research on this interview by examining the validity of the ChIPS with a larger sample of psychiatrically hospitalized adolescents than prior studies (i.e., Fristad et al., 1998a,b). The current findings are largely consistent with prior research demonstrating good concurrent validity in relation to other diagnostic interviews (i.e., the DICA-R-C; see Weller et al., 2000).

The percentage of agreement between the ChIPS and the K-SADS-PL in the present research ranged from 95% to 98%. The majority of the k coefficients (either standard or low base rate) for the 12 syndromes examined fell within the fair to moderate range. Specifically, agreement ranged from slight for ODD to substantial for substance use disorder. By demonstrating reasonable concordance between the ChIPS and the K-SADS-PL, a widely used diagnostic interview with demonstrated reliability and validity (e.g., Kaufman et al., 1997; see Ambrosini, 2000), these results add further support for the concurrent validity of the ChIPS.

More diagnoses of MDD, GAD, ADHD, and conduct disorder were made with the ChIPS compared to the K-SADS-PL. Only substance use disorder was endorsed more frequently by the K-SADS-PL than the ChIPS. Similar results have been observed in comparisons of the ChIPS to the DICA-R-C (e.g., Fristad et al., 1998a). The differing propensity for syndrome endorsements observed in the present research may be due in part to the different rating scales used in the instruments. The ChIPS scores symptoms as either present or absent. The K-SADS-PL differentiates whether a symptom meets threshold (i.e., is present) or subthreshold (i.e., is suspected/likely) or is absent. Thus, the ChIPS may be more likely to yield diagnoses considered subsyndromal by the K-SADS-PL (for a related discussion concerning agreement between the computerized DICA-R and the K-SADS-PL, see Hamilton and Gillham, 1999).

The ChIPS scoring format does not consider differential diagnosis at the time of the interview. In contrast, the K-SADS-PL specifically considers differential diagnosis at the time of syndrome endorsement. In the present study the impact of these differing criteria was most apparent when considering ChIPS endorsements of ODD. Specifically, 49 diagnoses of ODD were initially endorsed by the ChIPS compared to 15 ODD diagnoses endorsed by the K-SADS-PL. After taking into account comorbid MDD and conduct disorder, the number of ChIPS ODD diagnoses was reduced to 12. The ChIPS was designed to be a diagnostic screening instrument (Teare et al., 1998a; Weller et al., 2000); thus, some overascertainment is expected. In addition, it is likely that a clinically trained rater would have
<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-Level Differences (Δ) in Self-Reported Adjustment Indices for the</td>
</tr>
<tr>
<td>Depressive Disorder Composite, Anxiety Disorder Composite, Trauma</td>
</tr>
<tr>
<td>Composite, and Externalizing Composite by ChiPS Endorsements (i.e.,</td>
</tr>
<tr>
<td>Present, Absent) and by K-SADS-PL Endorsements</td>
</tr>
<tr>
<td><strong>Self-Report</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Depressive disorder composite (MDD and dysthymia)</td>
</tr>
<tr>
<td>RADS-2</td>
</tr>
<tr>
<td>BSS</td>
</tr>
<tr>
<td>SIQ</td>
</tr>
<tr>
<td>HSC</td>
</tr>
<tr>
<td>Anxiety disorder composite (GAD and social phobia)</td>
</tr>
<tr>
<td>MASC-10*</td>
</tr>
<tr>
<td>SCARED</td>
</tr>
<tr>
<td>Trauma composite (PTSD and acute stress disorder)</td>
</tr>
<tr>
<td>TSCC*</td>
</tr>
<tr>
<td>Externalizing composite (ADHD, ODD, and conduct disorder)</td>
</tr>
<tr>
<td>AQ</td>
</tr>
<tr>
<td>STAXI*</td>
</tr>
</tbody>
</table>

*Note: MDD = major depressive disorder; RADS-2 = Reynolds Adolescent Depression Scale; BSS = Beck Scale for Suicidal Ideation; SIQ = Suicide Ideation Questionnaire-Senior; HSC = Hopelessness Scale for Children; GAD = generalized anxiety disorder; MASC-10 = Multidimensional Anxiety Scale for Children-Short Version; SCARED = Screen for Child Anxiety Related Emotional Disorders; PTSD = posttraumatic stress disorder; TSCC = Trauma Symptom Checklist for Children; ADHD = attention-deficit/hyperactivity disorder; ODD = oppositional defiant disorder; AQ = Aggression Questionnaire; STAXI = State-Trait Anger Expression Inventory.

* T-scores are reported; all of the other scores are total scores.

*** p < .001.
VALIDATION OF THE CHIPS
implemented differential diagnosis considerations when interpreting syndrome endorsement by the ChIPS; however, the degree to which the interview proves to be overly inclusive could negatively affect the utility of the interview for both research and applied contexts.

We further examined the concurrent validity of the interview by comparing syndrome endorsement on the ChIPS with construct-specific self-report questionnaires. In each analysis adolescents diagnosed with a syndrome by the ChIPS scored significantly higher on self-report measures of impairment relative to adolescents who did not receive the diagnosis in question. In addition, the self-reported levels of distress and the magnitude of differences between those diagnosed with a ChIPS disorder compared to those who did not meet diagnostic criteria were nearly identical to parallel analyses conducted for the K-SADS-PL. Thus, the present study suggests that ChIPS diagnoses are largely consistent with self-report measures of impairment.

Cumulatively, the current multimethod approach suggests that the ChIPS is a valid instrument for identifying psychiatric illness and supports greater use of the ChIPS in clinical and research contexts. To achieve greatest utility and accuracy, a clinician should always be responsible for interpreting the results (Weller et al., 1999), ideally combined with a best-estimate clinical consensus procedure during which differential diagnoses are considered (e.g., Klein et al., 2001).

Limitations and Future Directions
Several limitations of this research should be noted. First, the effects of sex, age, and race on the validity of the ChIPS were not examined. The sample size, although significantly larger than prior research using the ChIPS with adolescents, precluded examining demographic differences in ChIPS- K-SADS-PL concordance. Relatedly, the results are limited to a small percentage of the potential participant pool (i.e., 20.7%, 97/46 met retention criteria) who were primarily European American (80%). Second, comparisons to treatment-team case consensus were unavailable. Thus, sensitivity and specificity analyses in relation to clinical case consensus could not be performed. Prior research indicates adequate sensitivity and specificity of the ChIPS in comparison to the DICA-R (e.g., Teare et al., 1998a; see Weller et al., 2000); future research would improve on this study by examining the sensitivity and specificity of the ChIPS relative to the K-SADS-PL. Such research would also provide evidence regarding whether the ChIPS is overly liberal and/or whether the K-SADS-PL is overly conservative. Third, mania/bipolar disorder and psychosis/schizophrenia were not addressed and should be in the future. Fourth, the present research relied exclusively on youth self-reports. Including parental informants would have been more consistent with the standard protocol for the K-SADS-PL (Kaufman et al., 1997). Obtaining information from other informants (e.g., parents) also would have improved confidence in the validity of diagnoses, especially those known to be underreported by children and adolescents (e.g., externalizing disorders [Loeber et al., 2000]). It is also possible that the observed relationships between the ChIPS and the K-SADS-PL and between the ChIPS and the questionnaires may have been enhanced because of reliance on a single informant (De Los Reyes and Kazdin, 2005). Fifth, in the present study the ChIPS, part of the standard intake of the inpatient facility from which the participants were drawn, was always conducted before the K-SADS-PL. It is important to note that the interviewers involved in the K-SADS-PL were blind to the results of the ChIPS; however, it is possible that the ordering of the interviews affected diagnostic endorsements. Finally, the K-SADS-PL was administered by one of six trained master’s degree–or postdoctoral degree–level clinical psychology trainees, whereas the ChIPS was administered by one of four highly trained bachelor’s degree–level assistants. Although the ChIPS was purposefully designed for administration by lay interviewers, differences in diagnostic endorsements made with the two instruments may in part be due to differences in the education level of the interviewers.

Clinical Implications
The ChIPS is a highly structured diagnostic interview that uses a branching format for assessing child and adolescent psychopathology with DSM-IV criteria (Weller et al., 2000). Advantages of the ChIPS include a relatively brief administration time, the use of simple language appropriate for use with children ages 6 to 18 years, and a highly structured format designed for lay interviewers (Weller et al., 2000). Studies conducted by the developers of the interview have provided initial evidence of the validity of the ChIPS in comparison to the DICA-R-C and to clinician diagnoses (e.g., Fristad et al., 1998a; see Weller et al., 2000).
The findings of the present research, in conjunction with the results of these prior studies, suggest that the ChIPS generates reliable psychiatric diagnoses among adolescents. These findings support greater use of the ChIPS in both clinical and research settings. Clinicians must be aware that the ChIPS may generate more diagnoses than the K-SADS-PL and should take care to follow differential diagnosis procedures outlined in the DSM-IV (American Psychiatric Association, 1994).

Disclosure: The authors have no financial relationships to disclose.

REFERENCES

Briere J (1996), Trauma Symptom Checklist for Children. Lutz, FL: Psychological Assessment Resources
Landis JR, Koch GG (1977), The measurement of observer agreement for categorical data. Biometrics 33: 159–174
Reynolds WM (1991), Suicidal Ideation Questionnaire. Odessa, FL: Psychological Assessment Resources
Sims J, Wright CC (2005), The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 85: 257–268
COPYRIGHT INFORMATION

TITLE: Validation of the Children’s Interview for Psychiatric Syndromes (ChIPS) With Psychiatrically Hospitalized Adolescents
SOURCE: J Am Acad Child Adolesc Psychiatry 46 no11 N 2007

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited.