Child Mania Rating Scale: Development, Reliability, and Validity

MANI N. PAVULURI, M.D., PH.D., DAVID B. HENRY, PH.D., BHARGAVI DEVINENI, M.D.,
JULIE A. CARBRAY, D.N.SC., AND BORIS BIRMAHER, M.D.

ABSTRACT

Objective: To develop a reliable and valid parent-report screening instrument for mania, based on *DSM-IV* symptoms.
Method: A 21-item Child Mania Rating Scale-Parent version (CMRS-P) was completed by parents of 150 children (42.3% female) ages 10.3 ± 2.9 years (healthy controls = 50; bipolar disorder = 50; attention-deficit/hyperactivity disorder [ADHD] = 50). The Washington University Schedule for Affective Disorders and Schizophrenia was used to determine *DSM-IV* diagnosis. The Young Mania Rating Scale, Schedule for Affective Disorders and Schizophrenia Mania Rating Scale, Child Behavior Checklist, and Child Depression Inventory were completed to estimate the construct validity of the measure. **Results:** Exploratory and confirmatory factor analysis of the CMRS-P indicated that the scale was unidimensional. The internal consistency and retest reliability were both 0.96. Convergence of the CMRS-P with the Washington University Schedule for Affective Disorders and Schizophrenia mania module, the Schedule for Affective Disorders and Schizophrenia Mania Rating Scale, and the Young Mania Rating Scale was excellent (.78–.83). The scale did not correlate as strongly with the Conners parent-rated ADHD scale, the Child Behavior Checklist -Attention Problems and Aggressive Behavior subscales, or the child self-report Child Depression Inventory (.29–.51). Criterion validity was demonstrated in analysis of receiver operating characteristics curves, which showed excellent sensitivity and specificity in differentiating children with mania from either healthy controls or children with ADHD (areas under the curve of .91 to .96). **Conclusion:** The CMRS-P is a promising parent-report scale that can be used in screening for pediatric mania. *J. Am. Acad. Child Adolesc. Psychiatry,* 2006;45(5):550–560. **Key Words:** child, bipolar disorder, rating scale, mania.

There is a growing consensus on the existence and description of the core symptoms of pediatric bipolar disorder (PBD; National Institute of Mental Health, 2002). PBD is a serious illness that can lead to high suicide rates, failure in school, aggression, and high-risk behaviors such as sexual promiscuity and substance abuse (Geller et al., 2002a, 2004; Pavuluri et al., 2005; Tondo et al., 1999). It is characterized by high relapse rates and low rates of recovery (Carlson and Kelly, 1998; Geller et al., 2004). Early recognition and accurate identification of PBD are necessary first steps toward prevention and early intervention (Zimmerman et al., 2004). Toward this goal, an instrument that is designed and tested specifically for pediatric mania is essential. In this article, we introduce a parent rating scale that is suitable as a screening instrument in clinical practice and research.

An important attribute of a valid screening instrument is the ability to distinguish PBD from attention-deficit/hyperactivity disorder (ADHD). These two

disorders share common features, such as hyperactivity, impulsivity, and distractibility (Geller et al., 2002a; Wozniak et al., 1995). A screening instrument must be able to differentiate PBD from ADHD (Dienes et al., 2002; Fristad et al., 1992).

Several rating scales are available for assessing PBD. Two of them are modified adult scales intended for parent ratings. The Parent Young Mania Rating Scale (P-YMRS) was validated by Gracious and colleagues (2002). This scale is a modified version of the YMRS, originally developed for hospitalized adult inpatients. Its content predates *DSM-IV* (American Psychiatric Association, 1994). The P-YMRS has reasonable psychometric properties, but it includes several items with poor factor loadings when rated by parents despite modifications for parent use (Gracious et al., 2002). The item content is not developmentally appropriate for children, for whom items asking about insight or appearance may be irrelevant (Axelson et al., 2004). A second promising measure is the General Behavior Inventory (GBI; Findling et al., 2002; Youngstrom et al., 2001). This is a 73-item adult rating scale that has both manic and depressive modules, with 6 additional "rule out" items. Apart from being lengthy, the test items are complicated and require at least a seventh-grade reading ability (Grammatik, Version 11.0, Corel Corp., Ottawa, Canada).

Youngstrom and colleagues (2004) reviewed the screening instruments for pediatric mania. They suggested that parent reports may have higher diagnostic accuracy than teacher reports or self-reports, and that these other reports add little information beyond parent reports. However, the areas under the curve (AUC) of parent-rated instruments tend to be modest (.78–.84). Thus, parents may be the most reliable source of information when screening for mania in children, but parent-rated instruments require improvement to use parent reports optimally.

Other parent-report instruments have been used to screen for PBD, but these instruments were not designed specifically for mania. For example, the Child Behavior Checklist (CBCL) has been used to provide markers for psychopathology, including mania (Biederman et al., 1995; Carlson and Kelly, 1998; Dienes et al., 2002; Geller et al., 1998). A consistent pattern of elevated scores was noted on Aggressive Behavior and Attention Problems, Delinquent Behavior, and Anxious/Depressed profiles (Kahana et al., 2003; Mick et al., 2003; Youngstrom and Youngstrom, 2005)

This congruence in pattern elicited by the CBCL may be caused by high comorbidity with ADHD, oppositional defiant disorder, conduct disorder, and anxiety disorders, and variable presentation of PBD children. However, the sensitivity of the CBCL for identifying mania was substantially lower than that of the mania-specific instruments cited above (Youngstrom and Youngstrom, 2005). Although low scores may be helpful in ruling out mania (or any psychopathology), high scores are not useful for ruling in mania (Youngstrom and Youngstrom, 2005). For this reason, a parent-report scale specifically designed to assess mania may be preferable to a more general scale such as the CBCL. The present study reports the development, reliability, and validity of such a symptom-specific scale, the Child Mania Rating Scale-Parent version (CMRS-P). A copy of this measure is available on the *Journal's* Web site at www.jaacap.com via the ArticlePlus feature.

METHOD

Subjects

The subjects for this study were 150 children (42.3% female) ages 10.3 ± 2.9 years. Thirty-one percent were 5 to 8 years old, 50% were between 9 and 12 years old, and 19% were 13 to 17 years old. The sample was evenly divided among healthy controls (HC), children diagnosed as having bipolar disorder, and children diagnosed as having ADHD. Subjects were included if they were between 5 and 17 years of age inclusive, had bipolar disorder I, II, or bipolar disorder-not otherwise specified (NOS); had ADHD; or were HC. Potential participants were excluded if they had head injury, epilepsy, or mental retardation, had significant medical illnesses, or were taking any medication or substance that could alter their moods. Patients that were receiving treatment in our pediatric mood disorders clinic were excluded because the purpose of the study was to screen for subjects at the intake phase, regardless of the severity of their disorders. To assess retest reliability and provide evidence for construct validity, we selected 20 subjects at random from the full sample. The parents of these children returned to the research center and completed the CMRS-P a second time, 1 week following their initial interviews. These parents were not informed of the purpose of completing the measure a second time to avoid reactive arrangements between the study design and the measure.

Figure 1 is a CONSORT chart detailing the recruitment of subjects for this study. Potential subjects believed at the outset to have bipolar disorder, ADHD or neither (HC) were recruited from various sources, such as our psychiatry intake clinic, community institutions/organizations, child psychiatrists, pediatricians, and the Child and Adolescent Bipolar Foundation. We recruited HC from community organizations (e.g., youth soccer teams). We initially assessed interest in "volunteering to help develop a measure to identify emotional problems," and then screened interested parents by telephone, inviting eligible participants to the research program for an interview. Each participant was paid
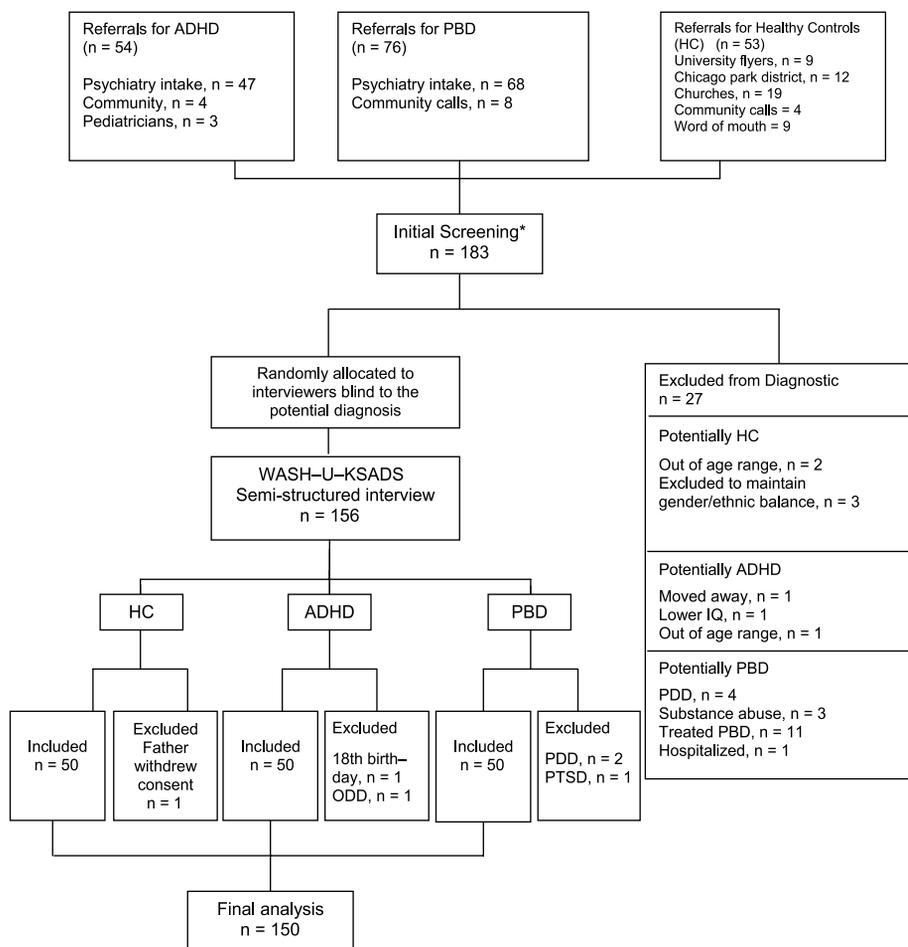
* Attempted to maintain similar male/female and ethnic compositions of potential diagnostic groups.

Fig. 1 CONSORT chart. Subject selection and retention. ADHD = Attention-deficit/hyperactivity disorder; PBD = pediatric bipolar disorder; ODD = oppositional defiant disorder; WASH-U-KSADS = Washington University Schedule for Affective Disorders and Schizophrenia; PDD = pervasive developmental disorder; PTSD = Posttraumatic stress disorder.

to cover the expenses for transport and parking charges at the time of the interviews.

At the outset of the study, we intended to match subjects in the three diagnostic groups on gender, ethnicity, age, and residence location. Subject availability and strict inclusion and exclusion criteria rendered this task extremely difficult to achieve in practice. Recognizing the difficulty of matching, we attempted to maintain reasonable gender and ethnic balance in the three screened groups by attending to these factors in recruiting. A total of 183 potential subjects were screened. Of these, 156 were assigned to be interviewed by one of three clinicians (the clinician was determined at random). The clinicians were blind to the screened diagnoses of the subjects they were assigned to interview. Semistructured diagnostic interviews using the Washington University Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS; Geller et al., 1998) were conducted to confirm the clinical diagnosis from the screening visit; Figure 1 details the results of this classification. The final sample consisted of 150 children, including

50 with bipolar disorder, comprising 34 with bipolar disorder I (68%), 8 with bipolar disorder II (16%), and 8 with bipolar disorder-NOS (16%); 50 with ADHD; and 50 HC. Of the 50 children with bipolar disorder, 42% had mixed mania.

Measures

Child Mania Rating Scale. The Child Mania Rating Scale, Parent Version (CMRS-P) is a mania rating scale designed to be completed by parents. In the present study, the measure was completed most frequently by mothers of the children (88%) and fathers, if they were the primary caregivers (12%). It includes 21 items reflecting the *DSM-IV* criteria for a manic episode (American Psychiatric Association, 1994). Each item is answered on a four-point Likert-type scale anchored by 0 (Never/Rare), 1 (Sometimes), 2 (Often), and 3 (Very Often; see the ArticlePlus feature on the *Journal's* Web site, www.jaacap.org). The entire scale is designed to be completed

in 10 to 15 minutes. In developing the CMRS-P, we took into account the following characteristics:

1. *DSM-IV basis.* Content-valid items on CMRS-P are designed to reflect symptom criteria A (levels of elation/irritability) and B (levels of self-esteem/grandiosity) for a manic episode according to *DSM-IV.*

2. *Singular item focus.* There are more items on the CMRS-P than there are *DSM-IV* criteria. For example, criteria A, consisting of elated mood or irritability under manic episode, is split into two items in the scale. Description of the independent items is an attempt to bestow clarity for parental use. Similarly, criteria B.1 (inflated self-esteem and grandiosity) is expanded into two items.

3. *Integrated functionality.* The *DSM-IV* criteria D on functionality is conserved in the instructions provided at the top of the rating scale. Each item is considered to be a problem only if it is causing trouble and is beyond what is normal for the child's age. Parents are instructed to check "never" or "rare" if the behavior is not causing trouble. Thus, the scale is designed to encourage parents to consider the possibility that their child's behavior is developmentally normative before labeling it pathological.

4. *Age-specific items.* Additional age-specific items are added toward the end. For example, the presence of rage attacks, and intense and prolonged temper tantrums can be strong associated features in younger age groups. Rapid mood swings are common in PBD. Furthermore, it is critical not to miss psychotic features that are not uncommon (Findling et al., 2001; Geller et al., 2002b). Therefore, screening questions are included to elicit these auxiliary features.

5. *Timing of symptoms.* It is pertinent to diagnose bipolar disorder if a cluster of symptoms co-occur at the same time period, instead of a single symptom manifesting at different time periods. If the total score of the rating scale is high enough, then it will reflect a cluster of symptoms that have occurred within the past month. Therefore, parents must take into consideration the child's behavior and emotions in the past month to synthesize the information required for screening to achieve a diagnosis of bipolar disorder. This time frame needs to be shortened to "past week" if this instrument is applied as an outcome measure to estimate weekly change during a clinical trial.

6. *Language.* We used the language familiar to and comfortable for parents, expanding the description to emphasize the meaning. Readability was tested using the Flesch-Kincaid reading level tests implemented through WordPerfect. The reading level of the entire measure was estimated to be at the sixth-grade level. We subsequently eliminated some complicated words and simplified sentences, reducing the reading level below sixth grade, a level similar to that recommended for informed consent forms by the Office of Human Research Protections (University of Illinois at Chicago). We also eliminated repetition of some critical words with each item (e.g., "periods of" as in "periods of elated mood") to make the instrument less tedious to complete. We included such expressions where necessary to communicate episodic manifestations of certain behaviors or emotions.

7. *Linked examples.* We included examples for several items in an attempt to enhance meaning and reduce ambiguity.

Other Mania Scales. To gather evidence for convergent validity, a research clinician evaluated each child using the WASH-U-KSADS, the YMRS, and the Schedule for Affective Disorders and Schizophrenia Mania Rating Scale (KSADS-MRS).

WASH-U-KSADS. This is a semistructured interview that is used to yield a *DSM-IV* diagnosis (Geller et al., 1998, 2001). This interview must be conducted by experienced clinicians. It includes an extensive list of items on affective symptoms, as well as a wide array of symptoms related to all other major psychiatric disorders. The WASH-U-KSADS also has a provision for documenting the onset and offset of rapid mood swings. Interrater reliability was 100% and the k value for mania and rapid cycling sections was 0.74 to 1.00.

YMRS. The YMRS is a clinician-rated measure (Young et al., 1978). It is not intended to be used as a screening tool, but rather to measure the symptoms in a manic state. The YMRS consists of 11 items, each with 5 explicitly defined levels of severity (0–4). The scale includes no depressive symptoms. Interrater reliability is 0.93 and concurrent validity with other mania rating scales is 0.77 to 0.89. The scale is designed to be administered by a trained clinician during a 15- to 30-minute-long interview. Severity ratings are based on the patients' subjective report of his or her condition during the past 48 hours and the clinician's observations during the interview. To compensate for uncooperative, severely ill patients, 4 of the 11 items—irritability, speech, content, disruptive-aggressive behavior—are given twice the weight of the other 7 items (i.e., rated as 0, 2, 4, 6, 8 rather than 0, 1, 2, 3, 4). This instrument was tested for use in children 5 to 17 years old (Youngstrom et al., 2002).

KSADS-MRS. KSADS-MRS allows for rating mania, including psychotic symptoms (Axelson et al., 2004). It is a clinician-rated scale with items similar to those on the KSADS structured interview (Orvaschel and Puig-Antich, 1987) and, therefore, can be scored from the interview if interview ratings are available. This scale consists of mania-specific, developmentally sensitive items. Decreased need for sleep (from WASH-U-KSADS) and mood lability are additional items not included on KSADS. The scoring is based on observations during the past week. Clinicians score the measure based on parent, child, and all other available information. Scoring ranges from 0 to 6 on each item (no information, not at all, slight, mild, moderate, severe, extreme). Scores on all of the items are added and 13 points are subtracted to yield a total score of 0 to 64. This instrument was tested in children ages 8 to 19 years (Axelson et al., 2004) and rated symptoms occurring in the past week. Internal consistency and interrater reliability were excellent at 0.94 and 0.97, respectively. Convergent validity with the Clinical Global Severity Scale was 0.91. The K-SADS-MRS scores differentiated bipolar patients who had clinically significant manic symptoms from those who did not, with a sensitivity of 87% and a specificity of 81%.

Other Psychopathology Measures. To estimate the construct validity of the CMRS-P, we asked parents to rate their children using the Conners Parent Rating Scale-Revised (CPRS-R; Conners, 1969; Conners et al., 1998) and the CBCL (Achenbach and Edelbrock, 1983). We also asked parents to complete a demographic questionnaire. Child or adolescent subjects completed a Children's Depression Inventory (CDI; Kovacs, 1985).

CPRS-R. CPRS-R is a 48-item parent-rating scale with 5 subscales: conduct problem, learning problem, psychosomatic, impulsive-hyperactive, and anxiety (Conners et al., 1998). In completing the scale, parents rate the presence and severity of symptoms in the child's current functioning. This scale was tested in 3- to 17-year-old children and covers the symptoms observed in the past month. Exploratory and confirmatory factor–analytic results of the CPRS-R revealed a seven-factor model including the following factors: cognitive problems, oppositional, hyperactivity-impulsivity, anxious-shy, perfectionism, social problems, and psychosomatic. The psychometric properties of the revised scale appear to be adequate as demonstrated by good internal reliability coefficients, high test-retest reliability, and effective discriminatory power.

CBCL. This version of the parent-rated behavior rating scale was validated for children between 4 and 16 years of age and covers the symptoms occurring in the past 6 months (Achenbach and Edelbrock, 1983). It is made up of eight factorially derived sub-scales: aggressive behavior, anxious-depressed, attention problems, delinquent behavior, social problems, somatic complaints, thought problems, and being withdrawn. Test-retest reliability was rated high, with moderately high agreement between parents, good content validity, and moderately high convergent validity.

CD. This measure was tested in 6- to 17-year-old children and asks questions about symptoms during the past 2 weeks (Kovacs, 1985). This measure comprises 27 items scored 0 to 2. It is a self-report instrument in which the child chooses one of the three sentences (normal through depressive) that describes him or her during the previous 2 weeks. Internal consistency is high (0.39–0.87). In terms of validity, it discriminates between depressed and nondepressed samples of children.

Procedure

The Institutional Review Board at the University of Illinois at Chicago approved the protocol used in the present study. All of the parents gave informed consent and children younger than 16 years old gave written or verbal assent, as required depending on the age of the child. For adolescents older than 16 years old, at least one parent and the adolescent gave written consent.

Diagnostic interviews were randomly assigned to interviewers, who were blind to the child's psychiatric diagnoses. The interviewers were one child psychiatrist (M.N.P.), one doctoral-level psychiatric nurse (J.A.C.), and two post-master's degree–level psychology graduates. Before conducting the interviews for this study, each had at least 6 months' experience administering the WASH-U-KSADS. Interrater reliability among the research interviewers at the end of training was .98 to 1.0 by Cohen's κ for the diagnosis on the WASH-U-KSADS. Individual-item reliabilities on the WASH-U-KSADS, YMRS and KSADS-MRS ranged from .92 to .96 by Cohen's κ. Monthly reliability checks were used throughout the study to maintain high interrater reliability. The research staff ($n = 6$) who administered the demographic and parent/self-report measures were different from those who conducted the diagnostic interviews ($n = 4$). The CMRS-P was completed by parents before conducting the diagnostic interview to minimize bias and fatigue effects.

Data Analysis

We evaluated the psychometric properties of the CMRS-P using classical test analysis, including measures of internal consistency, retest reliability, item–total correlation, confirmatory factor analysis, and correlational evidence for validity. To determine the optimal cut score for clinical classification, we calculated the sensitivity, specificity, and area under the receiver operating characteristics (ROC) curve for each possible cut score (Greiner et al., 2000; Swets, 1988).

We followed the construct validation approach described by Angold and Costello (2000) in gathering and presenting validity evidence for the CMRS-P. Validity analyses included the factor structure of the CMRS-P, the association between CMRS-P scores and scores on other instruments measuring pediatric mania, and the evidence that the CMRS-P can differentiate children with mania from HC and children with ADHD.

Logistic regression analysis provided evidence for the criterion validity of the CMRS-P and allowed us to choose optimal cut scores for use of the measure in clinical practice. From this analysis, we were able to evaluate the extent to which the CMRS-P accurately identified true positive and true negative cases (Tosteson and Begg, 1988). Sensitivity, specificity, and the AUC ROC varied depending on the cut score selected. We selected an optimal cutoff to differentiate mania from both ADHD symptoms and the difficult behavior found in healthy children.

We used SAS PROC LOGISTIC (SAS, Cary, NC) for comparing sensitivity, specificity and AUC for the three comparisons indicated above: children with bipolar disorder and children with ADHD, children with bipolar disorder and HC, and children with ADHD and HC.

RESULTS

Demographics and Comorbidity

Table 1 reports the demographic characteristics for the entire sample by diagnosis. There were no significant demographic differences among the three groups. Table 1 also reports the number of comorbid axis I diagnoses for the entire sample and for each diagnostic group. The bipolar disorder group had significantly more comorbid axis I diagnoses than either of the other groups (χ^2 [1, $N = 150$] = 121.2, $p < .001$). The most frequent comorbid diagnoses for the HC group were elimination disorder (10%) and adjustment disorder (2%). The most frequent comorbid diagnoses in the ADHD group were oppositional defiant disorder (32%), dysthymia (8%), and generalized anxiety disorder (6%). Among those in the bipolar disorder group, the most frequent comorbid diagnoses were ADHD (64%), oppositional defiant disorder (44%), and anxiety disorders (22%). Also, because the clinical diagnosis is the gold standard, it is important to clarify the line of thought behind what we consider to be the core symptoms. Deferring to the evidence from the psychopathology studies conducted in the PBD children (Geller et al., 1998), coupled with the centrality of mood features in bipolar disorder according to *DSM-IV* criteria, we considered elevated mood, grandiosity, or irritability as central features of PBD. These core clinical characteristics in our bipolar subgroup are summarized in Table 2.

Factor Structure

We used two methods to determine the unidimensionality of the CMRS-P. First, we examined the eigenvalues from an exploratory principal components analysis to determine how many of them would account for more variance than two items. This and the relative size of the eigenvalues of extracted factors can suggest the dimensionality of a matrix. Second, we tested the fit

TABLE 1
Demographics and Comorbidity

Variable	Total Sample	HC	Bipolar Disorder	ADHD	Test of Differences Between Samples (χ^2 or F)
Female sex, %	42.3	54.0	35.3	37.5	χ^2 (2) = 4.3, NS
Age, yr (mean ± SD)	10.3 ± 2.9	10.6 ± 0.4	10.4 ± 0.4	9.9 ± 0.4	$F_{(2,145)}$ < 1, NS
SES (mean ± SD)	1.8 ± 1.1	1.7 ± 1.2	1.8 ± 1.1	1.9 ± 0.9	$F_{(2,145)}$ = 1.2, NS
Race, %					
Caucasian	63.9	62.0	69.4	60.4	χ^2 (6) = 6.7, NS
AA	24.5	22.0	26.5	25.0	
Hispanic	10.2	16.0	2.0	12.5	
Asian	1.4	0.0	2.0	2.1	
Comorbid disorders					
ADHD, %	54.7	0	64	100	χ^2 (2) = 46.9, $p < .01$
ODD, %	26.7	4	44	32	χ^2 (2) = 15.8, $p < .01$
Anxiety disorders, %	10.7	2	22	8	χ^2 (2) = 9.8, $p < .01$

Note: SES = Mean revised Hollingshead socioeconomic status (Hollingshead, 1983); HC = healthy controls; ADHD = Attention-deficit/hyperactivity disorder; ODD = oppositional defiant disorder; AA = African American.

of a single-factor model using confirmatory factor analysis. Although four eigenvalues from an exploratory factor analysis were >1.0, only one was >2.0, indicating a factor accounting for more variance than two items. That eigenvalue was 11.73, accounting for 54% of the item variance. We tested the fit of a single-factor model using covariance structure modeling through LISREL-8 (Joreskog and Sorbom, 1993). This model provided an excellent fit to the data without the necessity of allowing correlated errors among items. (χ^2(189) = 61.24, NS; root mean square error of approximation = 0; adjusted goodness-of-fit index = .92). Table 3 reports the standardized factor loadings for each item from the confirmatory factor analysis.

Reliability

The internal consistency of the scale by Cronbach's α was .96 with the total sample. α Reliability coeffi-

cients within the diagnostic subsamples were .91 with both the bipolar disorder (n = 50) and HC (n = 50) samples and .92 with the ADHD (n = 50) sample. By age, the internal consistency of the CMRS-P was .97 for children ≤10 (n = 72) and .95 with children >10 (n = 78). The test-retest reliability of the measure was .96 at 1 week.

Content Validity

Figure 2 presents the percentage of the parent-rated sample responding that their children displayed symptoms "often" or "very often," by diagnoses of ADHD, bipolar disorder, or no diagnosis (HC). As shown in the figure, children with bipolar disorder had higher scores on most items than did children with ADHD or HC. On analyses of variance of each item with post hoc comparisons of raw scores, most of the items significantly differentiated the groups with a few exceptions.

TABLE 2
Core Clinical Characteristics of the BD Group (n = 50)*

	Elated Mood + Grandiosity + Irritability	Elated Mood + Grandiosity	Elated Mood + Irritability	Elated Mood	Irritability
Subjects, n (%)	28 (56)	3 (6)	15 (30)	1 (2)	3 (6)
BD I	24	3	6	1	0
BD II	4	0	3	0	1
BD-NOS	0	0	6	0	2
CMRS-P total score	36.82	32.67	27.13	15.00	9.00

Note: BD = Bipolar disorder; NOS = not otherwise specified; CMRS-P = Child Mania Rating Scale-Parent version.

*The groups were constructed based on the presence or absence of symptoms of elated mood, grandiosity, and irritability scored on the corresponding WASH-U-KSADS interview items.

TABLE 3
Factor Loadings From the Confirmatory Factor Analysis

Item	Loadings
Elated mood	.61
Irritable mood	.65
Grandiosity	.48
Delusional grandiosity	.56
Decreased need for sleep	.50
Too much energy	.64
Pressured speech	.63
Racing thoughts	.67
Flight of ideas	.61
Rushing around	.66
Distractibility	.62
Overproductive	.61
Hypersexuality	.51
Disinhibited	.56
Poor judgment	.58
Rage attacks	.62
Overly jocular	.58
Rapid mood swings	.63
Delusions	.61
Auditory hallucinations	.57
Visual hallucinations	.53

Note: All factor loadings were statistically significant at $p < .001$. Fit indices: Minimum fit function $\chi^2(189) = 61.2$, NS; root mean square error of approximation = 0; p(close) = 1.0; root mean square residual = .04; comparative fit index = .98.

The items "thinks that he/she can be anything or do anything," "believes he/she has unrealistic powers," "needs less sleep than usual," "rushes around non-stop," and "does many more things than usual/is unusually productive," did not differ between children with ADHD and children with bipolar disorder, but did discriminate between children with bipolar disorder and HC. Post hoc Bonferroni comparisons revealed that children with ADHD differed from HC only on the item "rushes around non-stop" ($p < .01$), but they did not differ on "thinks that he/she can be anything or do anything," "believes he/she has unrealistic powers," "needs less sleep than usual," or "does many more things than usual/is unusually productive."

Construct Validity

Evidence for construct validity is evidence of the extent to which the measure is measuring the construct it was designed to measure, and only that construct (Campbell and Fiske, 1959). We examined two types of such evidence. The first assessed the correlations between the CMRS-P and clinician-rating scales intended to measure the same construct (manic symptoms). For this purpose, three mania scales with established reliability and validity were used. The CMRS-P total score correlated .78 with the YMRS total score, .80 with the K-SADS MRS, and .83 with the WASH-U-KSADS Mania Module. The CMRS-P correlated moderately with the CBCL Aggression subscale (.51) and with the child self-report CDI (.45). We found modest correlations between the CMRS-P and the Conners ADHD scale ($r = .27$) and the CBCL Attention Problems subscale ($r = .29$).

Criterion-Related Validity

Table 4 reports the sensitivity, specificity, and AUC for these three comparisons. We found AUC of .91 and .96 for differentiating bipolar disorder from ADHD and bipolar disorder from HC respectively. A cut score of 20 resulted in a specificity of .94 and a sensitivity of .82. Tto avoid inclusion of the broad phenotype and reduce the possibility of overidentification, we excluded the eight children diagnosed with bipolar disorder-NOS from the ROC analysis. As would be expected, the measure did not discriminate well between ADHD and HC (AUC = .77). A cut score of 20 had poor sensitivity (.06) in this comparison. This finding underscores that the scale is specific for bipolar disorder, but it does not identify ADHD.

The positive likelihood ratio (LR+) is derived by dividing sensitivity by (1 − specificity) and it expresses the odds that a person scoring above a certain cut score actually has the disorder in question. As can be seen in Table 4, the LR+ for differentiating bipolar disorder from ADHD with a cut score of 20 is 13.67. This means that a person having a score of ≥20 on the CMRS-P is nearly 14 times as likely to have bipolar disorder as to have ADHD.

We also conducted an analysis of variance comparing the subtypes of the bipolar disorder diagnosis on the CMRS-P. The analysis found significant differences among the subtypes ($F_{(2,45)} = 6.38$, $p < .01$). Children diagnosed with bipolar disorder I (M = 35.8, S = 2.0) had significantly higher scores on the CMRS-P than did either children with bipolar disorder II (M = 23.4, S = 4.4) or bipolar disorder-NOS (M = 22.5, S = 4.1; all p values ≤ .05).

Finally, we compared CMRS-P scores on a fourfold diagnostic classification (HC, ADHD, bipolar disorder only, and BD+ADHD). Children with bipolar disorder
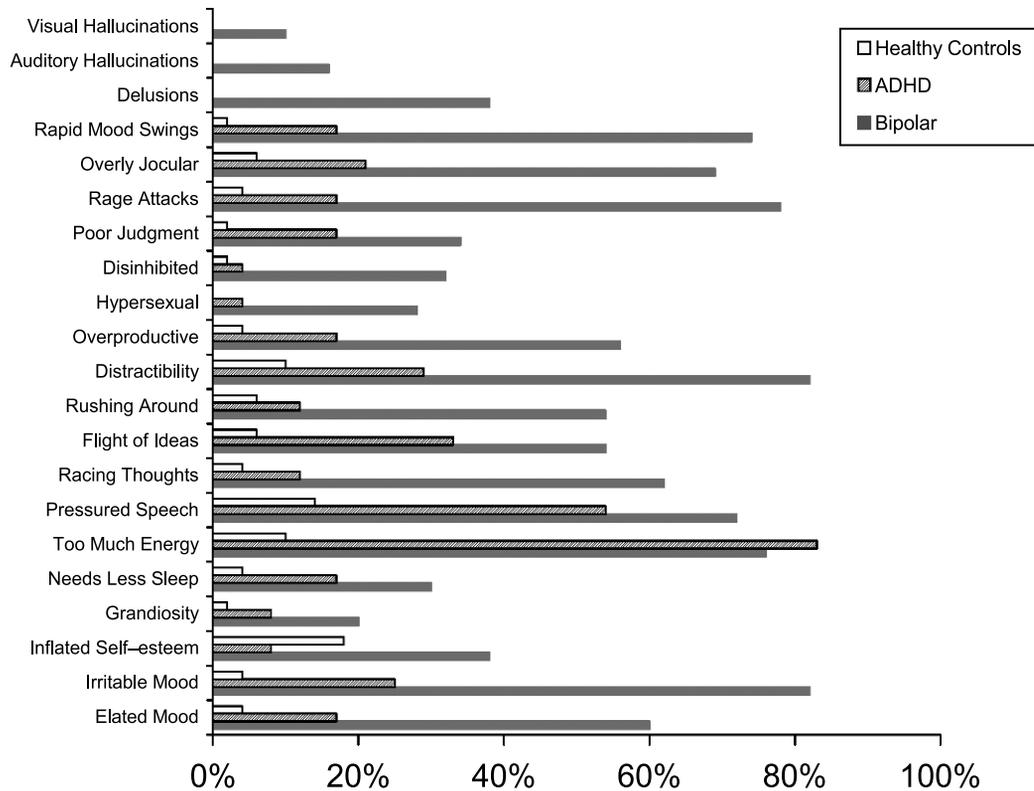
Fig. 2 Percentage of sample indicating presence of symptoms on Child Mania Rating Scale, by diagnosis (*n* = 150).

only and bipolar disorder + ADHD did not differ significantly in their scores (means of 27.31 and 33.41, respectively; *p* = .12), but children with either bipolar disorder only or bipolar disorder + ADHD had significantly higher scores than those with ADHD (mean of 9.94; *p* < .001) or HC (mean of 4.42; *p* < .001).

DISCUSSION

The CMRS-P is the first parent report measure developed specifically to assess child and adolescent mania. Internal consistency and retest reliability suggest that the CMRS-P is a reliable and valid instrument.

TABLE 4
Receiver Operating Characteristics at a Cut Score of 20

	AUC	Sensitivity	Specificity	LR+	LR-
BD vs. ADHD	.91	.82	.94	13.67	0.19
BD vs. HC	.96	.84	.98	42.0	0.16
ADHD vs. HC	.77	.06	.98	3.0	0.95

Note: BD = Bipolar disorder; ADHD = attention-deficit/hyperactivity disorder; HC = healthy controls; AUC = area under the curve; LR+ = positive likelihood ratio; LR- = negative likelihood ratio.

However, because of the small sample, the retest reliability findings should be interpreted with caution. Logistic analysis of the ROC curves suggest that the CMRS-P differentiates bipolar disorder from ADHD and HC, with high sensitivity and specificity, but, consistent with its purpose, does not differentiate ADHD from HC. Confirmatory factor analysis suggests that the CMRS-P is unidimensional. Evidence for the validity of the CMRS-P was obtained from strong correlations with the WASH-U-KSADS mania module, the KSADS-MRS, and the YMRS despite these measures being clinician-report scales.

In the present study, the ability of the CMRS-P to distinguish bipolar disorder from ADHD was shown in three types of analyses. First, the correlations between the CMRS-P and established ADHD measures (CPRS-R and the CBCL-Attention Problems subscale) were moderate. Second, our fourfold analysis (HC, ADHD only, bipolar disorder only, bipolar disorder + ADHD) found that the CMRS-P did not differentiate bipolar disorder + ADHD from bipolar disorder alone. Rather, it differentiates ADHD alone from bipolar disorder regardless of comorbid ADHD. Third, the ROC

analysis showed excellent sensitivity and specificity in differentiating children with mania from either HC or children with ADHD, indicating strong agreement with a reliable diagnostic interview that is regarded as a gold standard.

This greater ability to differentiate pediatric mania from ADHD or HC may be the result of steps taken in crafting the CMRS-P. Items that had low factor loadings on the P-YMRS such as those querying insight and appearance (Gracious et al., 2002) were excluded from the CMRS-P. Thought content, another item that had a poor factor loading on the P-YMRS, was subdivided into three CMRS-P items, inflated self-esteem, grandiosity, and any delusions. Our results indicate that a higher percentage of parents of children with ADHD endorsed the items indicating grandiosity and "needs less sleep," "rushes around non-stop," and "does many more things than usual/is unusually productive." It may be that parents of children with ADHD interpret high activity levels as equivalent to these items. Although this did not hamper the scale in differentiating pediatric mania from ADHD, it raises the question of parental interpretation and item responses. This underscores the importance of further clarification through diagnostic interviewing.

The CMRS-P adds to two other scales that have shown promise as parent-screening instruments for bipolar disorder in children. The P-YMRS (Gracious et al., 2002) has demonstrated good AUCs (.82) in differentiating bipolar disorder from unipolar depression and disruptive behavior disorders, but it has not been tested in its ability to differentiate bipolar disorder from HC. Although ruling out HC may be relatively easy with referred populations, it remains to be seen how a scale such as CMRS-P or P-YMRS performs in large epidemiological studies characterized by transient behavioral problems of varying degrees of severity. The General Behavior Inventory (Depression scale + Hypomanic/Biphasic scale), rated by parents, showed good AUCs (.88) and excellent sensitivity (.91) in screening for bipolar disorder, but it had low specificity (.68) in differentiating children with PBD from children with any other disorder or no disorder (Findling et al., 2002). Previous evidence for agreement among sources of information in screening for pediatric mania has not been strong (Findling et al., 2002), but this study provided evidence of high levels of agreement between parent and clinician ratings.

Despite its high sensitivity and specificity for screening pediatric mania, the CMRS-P correlated moderately with nonspecific instruments such as the CBCL and CDI. PBD presents with a high degree of irritability and symptoms associated with oppositional defiant disorder (Biederman et al., 1995; Findling et al., 2001, 2002; Geller et al., 2002a; Wozniak et al., 1995). These co-occurring symptoms may be responsible for the moderate correlation between the CBCL Aggressive Behavior subscale and the CMRS-P found in this study ($r = .51$). Similar findings have come from previous studies, including a meta-analysis of the CBCL scores from multiple data sets on the PBD (Mick et al., 2003), and comparisons of CBCL profiles across several urban settings and a community mental health center (Youngstrom et al., 2005). CBCL scores also indicated prominent depressive features in children with bipolar disorder (Mick et al., 2003) and mixed mania with depressive features is also common in PBD (Geller et al., 2002a; Pavuluri et al., 2005; Wozniak et al., 1995). It is also not surprising to see elevated scores on CDI with moderate correlation with CMRS-P.

The question of the breadth of the bipolar disorder phenotype has also been raised, suggesting that a narrower phenotype may be different from the broader phenotype (Leibenluft et al., 2003). Specificity in measures of core symptoms may help avoid overinclusion. In this study, the 34 children with bipolar disorder I had consistently higher scores than those diagnosed with either bipolar disorder II or bipolar disorder-NOS. In addition, we demonstrated that the CMRS-P was able to recognize the narrow phenotypes (bipolar disorder I and II; Leibenluft et al., 2003) by calculating the ROC excluding children with bipolar disorder-NOS.

This study suggests that the CMRS-P is useful in screening for mania. Future research will be needed to evaluate its sensitivity to change in symptoms during the course of treatment.

Limitations

One limitation is that this study was conducted with a sample enriched by referrals to a mood disorder clinic. Children may have already been diagnosed with bipolar disorder, and as a result, parents may have been well informed about bipolar disorder before being presented with this measure. There is no way to determine the extent to which such priming would have influenced the results. We took the following steps to minimize the

effects of such awareness. First, subjects participated in the research immediately after screening, and before beginning treatment. Second, we were careful to exclude children who were screened as having bipolar disorder, but were already under treatment at the time they were screened.

Another limitation is that, with an awareness of the controversy over the nature of bipolar disorder among children, we were careful to exclude most children presenting with irritability but without other associated symptoms of bipolar disorder. This may be a limitation to the extent that the measure with the cutoff score of 20 may not be as sensitive in identifying broader phenotypes as may be desirable.

The third and most important limitation is that this study was not conducted with an epidemiological sample. Because of this, the generalizability to community samples is not yet established.

Clinical Implications

An important implication for clinical use of this measure is that it should not be considered a diagnostic tool. The CMRS-P is designed to screen for current mania as opposed to delineation of subtypes of bipolar disorder, type and severity of episodes (depressive/manic), or comorbidity. A thorough diagnostic interview is essential to confirm and characterize this complex illness. However, because parents are often the first to suspect that a child has a serious psychiatric disorder, this study suggests that the CMRS-P may be a useful and accurate screening instrument for pediatric mania. The CMRS-P appears to be a potentially significant addition to the history of present illness, mental status, and family history in screening for pediatric mania and differentiating it from ADHD.

The CMRS-P is based on *DSM-IV* criteria for mania, employs a single focus for each item to minimize error in reporting, indicates the degree to which each symptom interferes with a child's functioning, and incorporates age-specific items applicable to an age range of 5 to 17 years. The CMRS-P also includes psychotic symptoms that could be missed unless queried. Finally, the scale is worded in such a way as to allow completion by parents, including those with limited reading ability. These advantages, coupled with clear examples integrated into many items, will facilitate early recognition and intervention to ameliorate this serious public health problem.

REFERENCES

Achenbach TM, Edelbrock C (1983), *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington: University of Vermont Department of Psychiatry

American Psychiatric Association (1994), *Diagnostic and Statistical Manual of Mental Disorder, 4th edition*. Washington, DC: American Psychiatric Association

Angold A, Costello E (2000), The Child and Adolescent Psychiatric Assessment (CAPA). *J Am Acad Child Adolesc Psychiatry* 39:39–48

Axelson D, Birmaher B, Brent D, Wassick S, Hoover C, Bridge J, Ryan N, (2004), A preliminary study of the Kiddie Schedule for Affective Disorders and Schizophrenia mania rating scale for children and adolescents. *J Child Adolesc Psychopharmacol* 13:463–470

Biederman J, Wozniak J, Kiely K, Ablon S, Faraone S, Mick E, Mundy E, Kraus I (1995), CBCL clinical scales discriminate prepubertal children with structured interview-derived diagnosis of mania from those with ADHD. *J Am Acad Child Adolesc Psychiatry* 34:464–471

Campbell DT, Fiske DW (1959), Convergent and discriminant validation by the multi-trait-multi-method matrix. *Psychol Bull* 56:81–105

Carlson G, Kelly K (1998), Manic symptoms in psychiatrically hospitalized children – what do they mean? *J Affect Disord* 51:123–135

Conners C (1969), A teacher rating scale for use in drug studies with children. *Am J Psychiatry* 126:884–888

Conners C, Sitarenios G, Parker J, Epstein J (1998), The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability and criterion validity. *J Abnorm Child Psychol* 26:257–268

Dienes K, Chang K, Blasey C, Adleman N, Steiner H (2002), Characerization of children of bipolar parents by parent report CBCL. *J Psychiatr Res* 36:337–345

Findling R, Gracious B, McNamara N, Youngstrom E, Demeter C, Branicky L, Calabrese J (2001), Rapid, continuous cycling and psychiatric co-morbidity in pediatric bipolar I disorder. *Bipolar Disord* 3:202–210

Findling R, Youngstrom E, Danielson C, DelPorto-Bedoya D, Papish-David R, Townsend L, Calabrese J (2002), Clinical decision-making using the General Behavior Inventory in juvenile bipolarity. *Bipolar Disord* 4:34–42

Fristad M, Weller E, Weller R (1992), The Mania Rating Scale: can it be used in children? A preliminary report. *J Am Acad Child Adolesc Psychiatry* 31:252–257

Geller B, Warner K, Williams M, Zimerman B (1998), Prepubertal and young adolescent bipolarity versus ADHD: assessment and validity using the WASH-U-KSADS, CBCL, and TRF. *J Affect Disord* 51:93–100

Geller B, Zimerman B, Williams M, Bolhofner K, Craney J, DelBello M, Soutullo C (2001), Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. *J Am Acad Child Adolesc Psychiatry* 40:450–455

Geller B, Zimerman B, Williams M, Delbello M, Bolhofner K, Craney J, Frazier J, Beringer L, Nickelsburg M (2002a), DSM-IV mania symptoms in a prepubertal and early adolescent bipolar disorder phenotype compared to attention-deficit hyperactive and normal controls. *Child Adolesc Psychopharmacol* 12:11–25

Geller B, Zimerman B, Williams M, Delbello MP, Frazier J, Beringer L (2002b), Phenomenology of prepubertal and early adolescent bipolar

disorder: examples of elated mood, grandiose behaviors, decreased need for sleep, racing thoughts and hypersexuality. *J Child Adolesc Psychopharmacol* 12:3–9

Geller B, Tillman R, Craney J Bolhofner K (2004), Four-year prospective outcome and natural history of mania in children with a prepubertal and early adolescent bipolar disorder phenotype. *Arch Gen Psychiatry* 61:459–467

Gracious BL, Youngstrom EA, Findling RL, Calabrese JR (2002), Discriminative validity of a parent version of the Young Mania Rating Scale. *J Am Acad Child Adolesc Psychiatry* 41:1350–1359

Greiner M, Pfeiffer D, Smith RD (2000), Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev Vet Med* 45:23–41

Hollingshead AB (1983), Tests and measures in social science. In: *Handbook of Research Design and Social Measurement, 4th ed.*, Miller DC, ed. Newbury Park, CA: Sage Publications

Joreskog K, Sorbom D (1993), *LISREL 8 User's Reference Guide.* Chicago, IL: Scientific Software International

Kahana SY, Youngstrom EA, Findling RL, Calabrese JR (2003), Employing parent, teacher, and youth self-report checklists in identifying pediatric bipolar spectrum disorders: an examination of diagnostic accuracy and clinical utility. *J Child Adolesc Psychopharmacol* 13:471–488

Kovacs M (1985), The Children's Depression Inventory (CDI). *Psychopharmacol Bull* 21:995–998

Leibenluft E, Charney D, Towbin K, Bhangoo R, Pine D (2003), Defining clinical phenotypes of juvenile mania. *Am J Psychiatry* 160:430–437

Mick E, Biederman J, Pandina G, Faraone SV (2003), A preliminary meta-analysis of the Child Behavior Checklist in pediatric bipolar disorder. *Biol Psychiatry* 53:1021–1027

National Institute of Mental Health Research 2002. Roundtable on bipolar disorder NOS in children. http://www.nimh.nih.gov/scientificmeetings/pediatric.cfm. Accessed February 12, 2005

Orvaschel H, Puig-Antich J (1987), *Schedule for Affective Disorder and Schizophrenia for School-Age Children: Epidemiologic, Fourth Version.* Ft. Lauderdale, FL: Nova University

Pavuluri MN, Birmaher B, Naylor M (2005), Pediatric bipolar disorder: 10-year review. *J Am Acad Child Adolesc Psychiatry* 44:846–871

Swets JA (1988), Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293

Tondo L, Baldessarini RJ, Hennen J (1999), Lithium and suicide risk in bipolar disorder. *Prim Psychiatry* 6:51–56

Tosteson A, Begg C (1988), A general regression methodology for ROC curve estimation. *Med Decis Making* 8:204–215

Wozniak J, Biederman J, Kiely K, Ablon J, Faraone S, Mundy E, Mennin D (1995), Mania-like symptoms suggestive of childhood-onset bipolar disorder in clinically referred children. *J Am Acad Child Adolesc Psychiatry* 34:867–876

Young R, Biggs JT, Ziegler VE, Meyer DA (1978), A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* 133:429–435

Youngstrom E, Youngstrom J (2005), Evidence based assessment of pediatric bipolar disorder, II: incorporating information from behavior checklists. *J Am Acad Child Adolesc Psychiatry* 44:823–828

Youngstrom E, Findling R, Danielson C, Calabrese J (2001), Discriminative validity of parent report of hypomanic and depressive symptoms on the General Behavior Inventory. *Psychol Assess* 13:267–276

Youngstrom E, Danielson C, Findling R, Gracious B, Calabrese J (2002), Abstract Factor structure of the Young Mania Rating Scale for use with youths ages 5 to 17 years. *J Clin Child Adolesc Psychol* 31:567–572

Youngstrom E, Findling R, Calabrese J, Gracious B, Demeter C, Bedoya D, Price M (2004), Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5 to 17 years. *J Am Acad Child Adolesc Psychiatry* 43:847–858

Youngstrom E, Youngstrom JK, Starr M (2005), Bipolar diagnoses in community mental health: Achenbach Child Behavior Checklist profiles and patters of comorbidity. *Biol Psychiatry* 58:569–575

Zimmerman M, Posternak M, Chelminski I, Solomon D (2004), Using questionnaires to screen for psychiatric disorders: a comment on a study of screening for bipolar disorder in the community. *J Clin Psychiatry* 65:605–610