Evidence-Based Assessment Strategies for Pediatric Bipolar Disorder

Eric A. Youngstrom, PhD,[1] Melissa McKeown Jenkins, MA, [1] Amanda Jensen-Doss, PhD, [2] and Jennifer Kogos Youngstrom, PhD[1]

[1] *Department of Psychology, University of North Carolina at Chapel Hill, North Carolina, U.S.A.*
[2] *Department of Psychology, University of Miami, Miami, Florida, U.S.A.*

ABSTRACT

Evidence-based assessment of pediatric bipolar disorder has advanced rapidly in the last two decades, moving from isolated clinical case descriptions to what is now a portfolio of techniques that include checklists from multiple informants, semi-structured diagnostic interviews and severity ratings, and technologies that allow daily tracking of mood and energy over the course of treatment. This review critically appraises (a) the need for evidence-based assessment of bipolar disorder as a common component of clinical practice, (b) triggers that warrant assessment of bipolar, (c) when best to deploy different techniques over the course of diagnosis and treatment, and (d) promising new developments in assessment. A decision-making framework is adapted from evidence-based medicine to guide assessment sequences in a patient-centered approach. Emphasis is placed on approaches that currently have the best validity and are feasible in most clinical practice settings. These methods increase accuracy and address many controversies surrounding pediatric bipolar diagnoses.

Conventional wisdom was that bipolar disorder most often manifested during young adulthood. Although there were occasional case reports in childhood or early adolescence, pediatric bipolar disorder (PBD) was considered exotic, and it was not part of the core training for physicians or mental health professionals working with youths. Even now, most textbooks and training materials focus on bipolar disorder as an "adult" condition. As a result, practitioners have had minimal training in the assessment of PBD. Should busy clinicians invest the time and effort to learn about evidence-based assessment strategies for pediatric bipolar disorder? Given the stakes involved in making this diagnosis correctly, as well as the rapid advances in the evidence base over the last several years, there are few niches that could provide so substantial a return on investment. Other papers in this special issue review the distinction between PBD and other forms of mood dysregulation and aggression (1, 2) and the evidence for the validity of PBD as distinct from ADHD, depression, or other more common developmental psychopathology (3). This review will address key topics, such as why to assess for PBD, when it is clinically indicated, how to change assessment strategies to match the individual needs of the patient over the course of treatment, and what promising future directions might be worth adding to clinical practices in the future.

WHY ADD FORMAL ASSESSMENT PROCEDURES FOR PBD TO THE CLINICAL TOOLBOX?

The place where PBD seems most scarce is in textbooks. There are now several thousand peer reviewed articles describing and validating pediatric bipolar disorder, drawn from dozens of independent research groups around the world (4). A recent meta-analysis of epidemiological studies found that ~2% of children and adolescents in the community - not clinics - meet criteria for bipolar spectrum diagnoses (5), with equal rates in the U.S.A. versus the rest of the world. A Canadian study published after the meta-analysis replicated the 2% figure (6). Increased rates of clinical diagnoses started from a baseline of almost never diagnosing PBD (7), and clinical and epidemiological rates are now converging.

Where have these bipolar cases been hiding? Often in

Address for Correspondence: ✉ Eric Youngstrom, Department of Psychology, University of North Carolina, CB #3270, Davie Hall, Chapel Hill, NC 27599-3270, U.S.A. eay@unc.edu

plain sight. Both community (8, 9) and clinical studies (10-14) indicate that PBD is highly impairing. However, when families seek services, PBD often is missed. If the mood symptoms are prominent, then the most likely diagnosis is major depression, contributing to the finding that one third of all cases with depression prove to have a bipolar spectrum disorder when followed longitudinally (15, 16). If the energy and attention problems are salient, then the likely diagnosis is ADHD or ODD, particularly in Europe and Israel, whereas aggression is more likely to attract a conduct disorder label, and psychosis a schizophrenia diagnosis – especially in ethnic minorities in the U.S.A (17). Because bipolar is an episodic illness, with dramatically different presentations during different phases, it is exceptionally challenging for clinicians to develop a good prototype for the "typical" case. Prototype matching is a main way that experienced clinicians formulate cases (18), but it performs badly with PBD (19, 20). Vignette studies demonstrate tremendous range of opinion, often varying by global region, when clinicians examine cases with potential PBD (21, 22). Clinical PBD diagnoses rarely agree with each other or with structured interview results at better than chance rates (23), contributing to the long lag between onset of problems and diagnosis in youths (24) and adults (25, 26).

Diagnostic disagreement is less surprising considering the dearth of formal training about PBD, forcing practitioners to learn as they work. Using evidence-based assessment tools can help close the gap, especially if practitioners can easily incorporate the methods without additional time, expense or training. PBD assessment has evolved rapidly, with dozens of tests now having published validity data. The challenges now are choosing the best from among contenders, and understanding each tool's role at different stages of assessment and treatment.

WHEN IS ASSESSMENT OF PBD CLINICALLY INDICATED?

Evidence-based medicine (EBM) uses probability of diagnosis as a way of organizing clinical decisions about assessment and treatment (27). Every case has a possibility of having PBD, albeit often low. Test scores, risk factors, and other pieces of evidence refine our estimates of the probability. When the probability is sufficiently low, the diagnosis can be "ruled out," at least until new information triggers re-evaluation (see Figure 1). When enough confirmatory evidence accumulates, then the probability rises enough that we make the diagnosis and concentrate on organizing the treatment around it. This Bayesian framework is similar to clinical thinking such as the "Bipolarity Index of Suspicion" (28, 29). Figure 1 illustrates how diagnostic probability maps onto clinical actions. EBM refers to two major choice points along this continuum: The Wait-Test Threshold, and the Test-Treatment Threshold (27). Below the Wait-Test Threshold, a diagnosis is considered "ruled out." Above the Test-Treatment Threshold, a diagnosis is considered firm enough to begin treatment. Between the two thresholds is where additional assessment is needed to either push the probability below the Wait-Test or above the Test-Treatment Threshold. EBM does not attach specific numbers to these thresholds. Where to set the bar is a clinical judgment, depending on risks and benefits. Formal approaches for integrating these utilities into decision-making may gain popularity as technology reduces the inconvenience associated with computation (27, 30). Table 1 lists steps in evidence-based assessment of PBD detailed below.

AN EVIDENCE-BASED PROCESS FOR THE DIAGNOSIS OF PBD

- **Step 1.** Know the base rate of PBD in your setting. The first piece of evidence to incorporate in fast, frugal PBD assessment is its base rate in a clinical setting. PBD rates vary widely depending on where one works. PBD is rare in the general community, but somewhat more common in outpatient practices, and even more in practices that specialize in mood disorders. Table 2 lists benchmarks from different settings. In many settings, PBD rates will fall below the clinician's Wait-Test threshold. For example, if a clinician decides that conditions seen in fewer than 1 in 20 cases do not warrant extra assessment unless other warning signs are evident, then their Wait-Test threshold is 5%. If working where <5% of cases might have PBD, then they do not need to include PBD assessment measures as part of their standard intake procedure. When the target is already rare, low scores on the test will not add information, and high scores will still usually be false positives. On the other hand, if working where PBD might be more common - such as an inpatient unit - assessment methods can quickly move some cases below the Wait-Test threshold, and others closer to the Treatment zone.
- **Step 2.** Assess PBD risk factors. There are risk factors and cues that should trigger further assessment. Most well-established is a family history of bipolar disorder (31, 32) (see Table 3). Bipolar in a first degree relative is linked with at least a five-fold increase in risk for PBD, and second-degree relatives with at least half as much risk (33). Other warning signs for PBD

Figure 1. *Integrated Model of Evidenced-Based Assessment for Pediatric Bipolar*

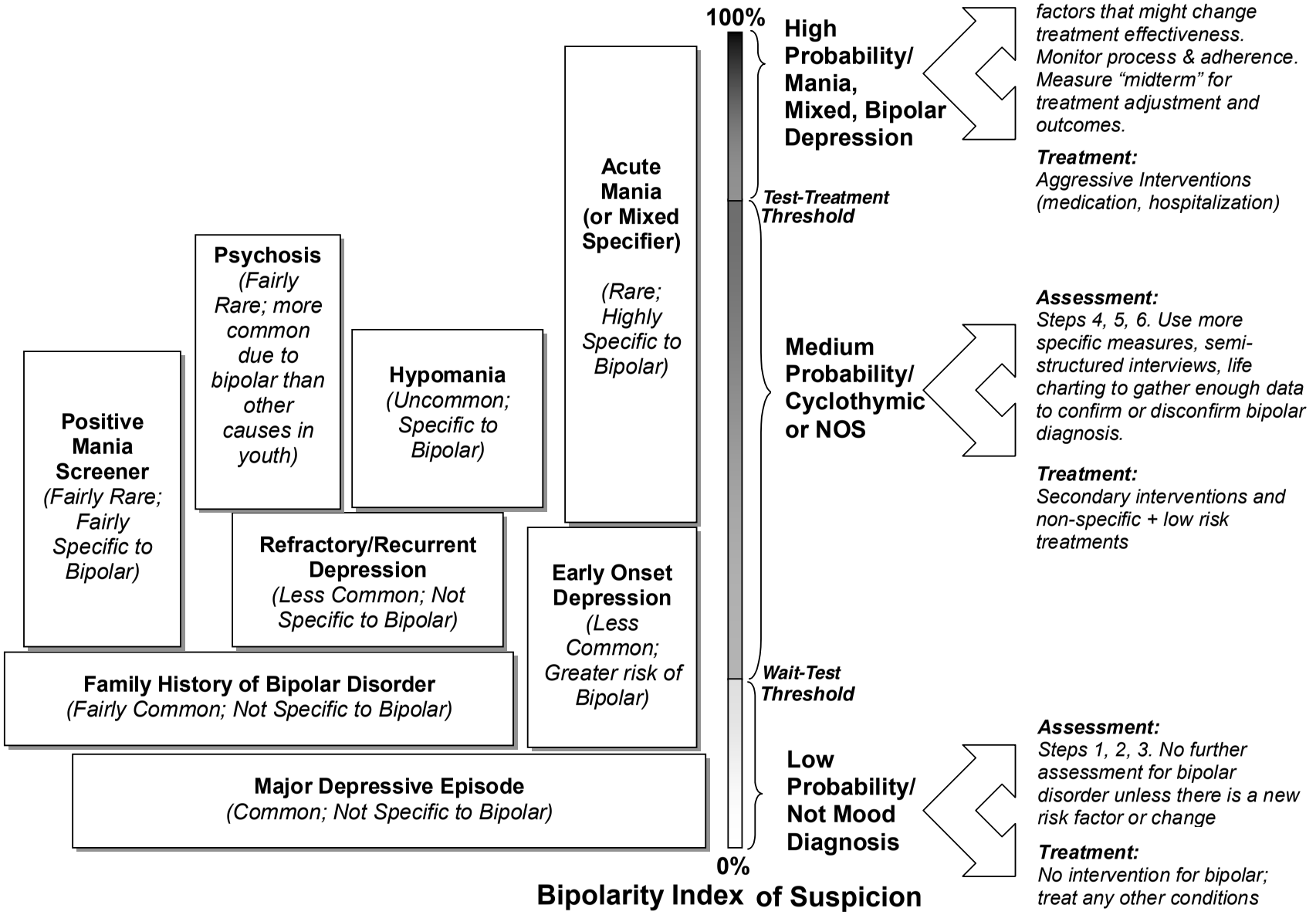


Table 1. *Ten Steps of Evidence-Based Assessment for Pediatric Bipolar Disorder*

Step	Rationale	Additional Time and Cost
Know base rate in your setting	Important starting point to anchor evaluations	Time: 0 Cost: 0
Any risk factors?	Risk factors raise "index of suspicion," enough in combination will elevate into assessment or possibly treatment zones	Time: 2-10 min Cost: 0
Information from broad, externalizing scales	Low externalizing on parent report usually rules bipolar out High parent report another "red flag" High youth report, teacher report double odds; Low scores less informative	Time: none if already part of routine assess Cost: : none if already part of routine assess
Add brief screens for family history, hypomania, mania	Brief family history measure may add new information Parent report screens replace Externalizing score – more specific to bipolar, but highly correlated (no "double dipping")	Time: 5 minutes for family, 2 minutes for practitioner Cost: None – best instruments are in public domain
Get multiple perspectives – and plan for differences	Parent report helpful in establishing diagnosis, change in functioning; youth and teacher report helpful for measuring pervasiveness and also motivation for treatment	Time: 5 minutes for each informant, 2 minutes for practitioner Cost: None – best instruments are in public domain
Intensive Assessment for bipolar	Clinical interview focusing on mood presentation and specific symptoms Semistructured interviews: KSADS, MINI Life charting – paper, online, smartphone application	Time: 30-120 minutes Cost: 0 to US $4.00 for applications (ILS 0 to 15 new shekels)
Additional assessment for treatment planning	Rule out general medical conditions, other medications; Family functioning, quality of life, personality, school adjustment, comorbidities	Time: Variable Cost: Variable
Process monitoring ("quizzes and homework")	Life charts, mood & energy checkups at each visit, medication monitoring, therapy assignments	Time: < 5 min per day for family, < 5 min per visit for practitioner Cost: None
Progress and outcome ("midterm and final exams")	Repeat assessment with main severity measures – interview and/or parent report most sensitive to treatment effects	Time: 10 to 40 minutes Cost: None
Maintenance	Discuss continued life charting; review triggers, critical events and life transitions	Time: Negligible Cost: None

Table 2. *Base rates of pediatric bipolar disorder in different settings.*

Setting	Base Rate	Population
General Population	2% (4% for spectrum)	Global meta-analysis (5)
Outpatient or Community Mental Health	5% to 10%	Various
County Wards (DCFS) (93)	11%	State of Illinois
Specialty Outpatient Service (94)	15% to 20%	New England, Midwestern USA
Incarcerated Adolescents (95) (96)	2% to 22%	Chicago & Texas
Inpatient and Psychiatric Hospitalization (97) (50)	25 to 40%	All of U.S.A. (record surveillance)

include psychosis – more commonly due to a mood disorder than schizophrenia among children or adolescents (34); early onset depression – which appears to be bipolar spectrum illness in a third of cases during follow-up (15); and sleep disturbance – especially periods of decreased need for sleep without associated fatigue (35). Additional clinical presentations that warrant increased attention include bouts of *episodic* aggression (2, 36) or someone initiating a referral specifically to evaluate PBD. Although these frequently prove to have a different etiology, PBD should be discounted based on disconfirming evidence.

- **Step 3.** Evaluate information from broad measures. Many clinicians routinely use instruments measuring multiple factors (37, 38). These types of instruments are sensitive to PBD (meaning that most cases with PBD score high), but not very specific to PBD (meaning that non-bipolar cases also tend to score high on the same scales) (39-41). Although PBD frequently involves a "profile" of elevations on multiple problem behavior scales, the Externalizing score captures most information relevant to possible bipolarity: If Externalizing is extremely elevated, the odds of PBD triple or quadru-

Table 3. *Clinical "red flags" that should trigger thorough evaluation of possible pediatric bipolar disorder*

Red Flag	Description	Reason
Family history of Bipolar*	PBD has genetic contribution Family environment can amplify risk Family environment affects treatment adherence and relapse	5x – 10x increase for 1st degree relative 2.5x-5x for 2nd degree relative 2x for "fuzzy" BP in relative Probe histories of depression, suicide, alcohol and drug, psychosis, and antisocial behavior for possible undiagnosed bipolar (33, 98)
Early Onset Depression	Onset < 25 years Also treatment resistant, recurrent, or atypical depression may be more likely to be bipolar	First clinical episode is often depression 20% to 35% of pediatric depressions ultimately show bipolar course (99, 100)
Antidepressant Coincident Mania	Manic symptoms while being treated with antidepressants	FDA recommends assessing for hypomania, family history of bipolar before beginning antidepressant "Switch" is often previously undiagnosed PBD (101)
Episodic Mood Lability	Rapid switching between depressive and manic symptoms; depressive and manic symptoms at the same time	Common presentation Episodicity more suggestive of mood diagnosis (4)
Episodic Aggressive behavior	Episodic; high-energy. Not instrumental or planned; reactive	Not specific, but common (4, 36)
Psychotic features	True delusions/hallucinations in the context of mood	Delusions/Hallucinations common during mood episode Bipolar more common as source of psychosis than schizophrenia in children (34, 36)
Sleep Disturbance	Decreased need for sleep Less sleep but maintains high energy	More specific to bipolar Indicates sleep hygiene component of treatment

ple. Conversely, low Externalizing scores make PBD much less likely, decreasing the odds by a factor of as much as 20 (40). Externalizing scores provide a distilled "bottom line": Although PBD frequently shows a constellation of multiple elevations across subscales, no other scales provide significant incremental information after interpreting Externalizing (40, 42, 43).

These initial steps in assessment (i.e., considering the base rate, risk factors, and scores on broad tests) add little or no time to the first interview. The steps reorganize readily available data according to value with regard to PBD. Patients with two or three of these "red flags" are too risky to ignore possible PBD, but also often will not have PBD. These are the cases where more assessment is clearly indicated.

- **Step 4.** Consider adding a mania measure. The fourth step in assessment would be to gather brief screening instruments focusing on manic symptoms, ideally from the parent or adult most familiar with the youth's behavior. Some manic symptoms will be more specific to PBD even if they are not associated with the most impairment. Symptoms such as elated mood, grandiosity and unstable self-esteem, and decreased need for sleep have fewer likely causes than irritability and aggression. Instruments that include more content specific to mania (44-46) significantly outperform broad scales at discriminating PBD from non-PBD. It is easiest for clinicians to simply substitute them for externalizing scores when formulating impressions about possible bipolarity. The more specific mania checklist replaces the Externalizing score in PBD algorithms, because it is the single most valid checklist score capturing the parent's impressions about the key behaviors for diagnosis (47).

  Because mania scales are more specific to PBD, high scores help move towards ruling the diagnosis "in." EBM has a mnemonic, "SpPIn and SnNOut": On a Specific test, a Positive score helps rule In; on a Sensitive tests, a Negative score helps rule Out (27). Low mania scale scores may cancel out one or two risk factors, and high scores provide a stronger surge in the index of suspicion. Note the asymmetry: Finding a family member affected with bipolar increases the PBD probability, but lack of a family history does not decrease risk as much (47). Low scores on sensitive tests (e.g., CBCL Externalizing) are more decisive than high scores on the same instrument (27). Table 4 compares the CBCL and measures more specific to PBD.
- **Step 5.** Get multiple perspectives – and plan for differences in view. In every published study to date, parent-reported manic symptoms consistently show greater diagnostic validity than youth- or teacher-report (40). Parent report shows large effect sizes discriminating PBD from non-PBD cases, whereas youth report

Table 4. *Areas Under the Curve (AUCs) and Likelihood Ratios for Selected Measures When Discriminating PBD from All Other Cases Seeking Outpatient Mental Health Services*

Screening Measure	AUC (Citation)	Test Score	Diagnostic Likelihood Ratio
Adolescents (11 to 18)			
Parent General Behavior Inventory -- Hypomanic/ Biphasic (46)	.84 (40)	<15 16-24 25-39 40-48 49+	.2 1.1 2.2 4.8 9.2
CBCL Externalizing *T*-Score (102)	.78 (40)	<54 54-56 65-69 70-75 76-80 81+	.04 .5 1.3 2.1 2.7 4.3
YSR Externalizing *T*-Score (103)	.71 (40)	<49 49-55 56-69 70-76 77+	.3 .5 1.4 2.3 3.0
Adolescent General Behavior Inventory – Hypomanic/Biphasic (104)	.64 (40, 41)	<10 10-37 38-45 46+	.3 1.0 2.0 3.9
Children (5 to 10 years)			
CBCL Externalizing *T*-Score (102)	.82 (40)	<58 58-67 68-72 73+	.1 .5 1.5 3.9
Parent General Behavior Inventory -- Hypomanic/ Biphasic (46)	.81 (40)	<11 11-20 21-30 31-42 43-50 51+	.1 .5 1.3 2.3 4.9 6.3

Note: All studies used some version of KSADS interview by a trained rater combined with review by a clinician to establish consensus diagnosis. Diagnostic Likelihood Ratio (DLR) refers to the change in probability associated with the test score. Likelihood ratios of 1 indicate that the test result did not change impressions. DLRs larger than 10 or smaller than .10 are frequently clinically decisive; 5 or .2 are helpful, and between 2.0 and .5 are small enough that they rarely result in clinically meaningful changes of formulation (27).

shows medium effects, and teacher report falls in the medium to small range, often not significant (48, 49). When multiple informants note problems, the youth has a more severe and impairing condition (50, 51). When youths have PBD, then their self-reported or teacher-reported levels of behavior problems are significantly higher than typical for youths with other diagnoses (52). However, cross-informant agreement about youth functioning is modest, commonly hovering in the range of $r = .2$ to .3 (53), so "statistically significant" agreement frequently looks like contradictory perspectives at the level of the individual case. For example, if the parent reports an Externalizing score of 80, then a teacher reporting a score of 60 looks like a major difference of opinion unless benchmarked against typical agreement levels (teacher scores would average 56 for that level of parental concern) (52).

Thus seeming disagreement may actually reflect a fair amount of cross-situational consistency, and one informant is often much more concerned about mood symptoms than others. Contrary to conventional belief that youths are the most accurate informant about their own moods (54), parent report is more accurate for the purpose of detecting PBD, possibly because hypomanic symptoms are not dystonic to the person experiencing them and because compromised insight is a feature of hypomania and mania (55, 56). Teachers do not observe some hallmark features of mania, such as decreased need for sleep; and they also tend to attribute many of the other symptoms to ADHD or oppositionality (57). Clinical judgments about individual credibility have meaningful impact on the reliability and validity of parent and youth report. However, parent credibility on average was not connected to their current stress level or history of mood disorder (58). Similarly, parent report remained significantly more valid even if the parent has a history of mood disorder. The evidence-based plan of action would be (a) always try to involve a parent in the evaluation of potential PBD, (b) consider multiple reporters as informative about the degree and contexts of impairment, but less helpful in differential diagnosis (47), and (c) use clinical judgment to decide about specific people's credibility, rather than using simple heuristics (such as "discount parents who have bipolar disorder themselves") that have failed to demonstrate statistical validity (58).

- **Step 6.** Intensive assessment methods for PBD. Through application of the first five steps, practitioners will be able to rule PBD *out* in the majority of cases, and with a high degree of confidence. In statistical parlance, cases testing "not bipolar" will have a high negative predictive value – these decisions will almost always be correct in most settings. This is valuable information: When bipolar is ruled out, something else will usually be ruled in. Once bipolar has been excluded, then the practitioner can treat that other condition with greater confidence. If the diagnosis is ADHD, then stimulant medication can be tried, or antidepressants used for depression or anxiety, with less concern about possible adverse events (59). The

practitioner will also be able to document the "due diligence" about considering alternative diagnoses

It is the "test positives"- the cases with a family history and/or high scores on Externalizing, and then a high score on a mania screen - that are more ambiguous. The index of suspicion/probability estimate will be in the mid-range for these cases. So long as we do not lump probabilities of 51% or higher together and treat them as "bipolar," we will not be "over-diagnosing." Mid-range cases are the ones where systematic, intensive assessment of bipolarity is indicated.

At a minimum, more intensive assessment of PBD involves a thorough clinical evaluation, combining interview with the youth, direct observation of mental status, and discussion with at least one collateral informant most familiar with the youth's behavior. Structured approaches (60-62) cover all symptoms of hypomania and mania, even if the family does not see them as a central part of the presenting problem. Clinicians are often reluctant to use structured approaches, believing that clients dislike structured approaches, or that they damage rapport, are not reimbursable, or they constrain professional autonomy (18). Surprisingly, none of these concerns are supported by data: client surveys indicate they prefer the thoroughness of structured approaches, without decrement in rapport or engagement (63). Medicaid and insurance companies will reimburse for intensive diagnostic interviewing, especially when screenings or other findings establish medical necessity. Semi-structured approaches also offer more latitude for clinical judgment while retaining the comprehensiveness of a structured approach (61, 62). Agreement between "clinical diagnoses as usual" and semi-structured approaches is typically poor (64), particularly for bipolar disorder (23, 65). Semi-structured approaches are more reliable, detect more comorbidity, and follow DSM criteria more consistently – reducing differences due to training or conceptualization (22, 66) and also shrinking potential bias due to race or ethnicity (67). We propose a hybrid model: if regular intake does not use a semi-structured interview, then the practitioner can add the mood modules of an interview such as a KSADS whenever the index of suspicion is in the "Assessment Zone." Additional diagnostic modules could also be selected to investigate competing diagnostic hypotheses or common comorbidities.

Other methods are also available for helping establish a PBD diagnosis. One of the most promising is "life charting" or "mood charting" (68), tracking changes in mood and energy on a daily basis. The diagnostic power emerges from the repetition and finer-grained resolution compared to retrospective reports about several weeks or a lifetime of functioning. Life charting evolved rapidly from paper and pencil measures to electronic versions that are available on the Web or as applications for smart phones (a Google search for "mood charting" finds the most current versions). The decrease in cost and advances in convenience and functionality make these an attractive method for gathering data about mood and energy changes. These are valuable in documenting pronounced shifts in energy and affect, and they also become an aid in monitor progress during treatment, as detailed below.

Other assessment techniques have accrued some research or clinical interest, but cannot be considered "evidence-based assessment" tools for clinical application yet. These include genetic tests, fMRI and other imaging tests, neurocognitive assessments, projective testing such as the Rorschach, or even published instruments that are commercially distributed but do not have any peer reviewed studies demonstrating validity. Imaging and genetics are only beginning to use clinically more realistic designs, with fewer exclusionary criteria and a greater emphasis on generalizability, including high rates of depression or ADHD in the comparison group (69).

- **Step 7.** Assessment for treatment planning. After gathering enough information to make a clear decision about PBD status, the last step before shifting to acute treatment is to collect any other key information that might change our choice of treatment: common comorbidities (36), current medications or substance use (which might interact with other drugs), prior medication trials, personality traits, stability of the family, academic functioning, and quality of life (70). Comorbid substance use may change the initial approach to treatment. Conversely, comorbid ADHD may not require different intervention initially, but instead involves monitoring whether symptoms subside with successful mood stabilization, versus warranting adjunctive intervention after mood stabilization has had time (59). Careful review of medical history ensures that what appear to be mood symptoms are not the result of some other general medical condition or drug side effect. Bipolar disorder is linked with low levels of personality traits such as conscientiousness, which often lead to forgetting appointments, home-

work, medication, and other accidents of omission that can undermine treatment (71). Poor family functioning predicts earlier onset, poor response during treatment, and rapid relapse (59). Effective interventions that reduce conflict and improve communication are being explored as adjunctive treatments for PBD (59). Short and simple instruments assessing family functioning may be most practical for busy practitioners. Assessing quality of life can help define objectives beyond mere symptom reduction and thus improve engagement with treatment. Again, some of the best tools are short and also in the public domain (72). Finally, it is crucial to assess and document potential suicidality (73).

- **Steps 8 & 9.** Measuring process and outcome. Once sufficient data confirm a PBD diagnosis, the focus shifts to treatment (27). Assessment plays a different important role during treatment (74). The key questions change from "what is the problem?" and "what are the causes?" to "how bad is the problem?" and "are we making progress towards our goals?" The simple acts of measuring severity and defining targets are themselves associated with better outcomes (75). The idea of going on a diet without measuring weight is absurd. It is similarly helpful to mark the starting point and goal while trying to manage mood and behavior. Some methods used in differential diagnosis will not contribute as much during active treatment. Repeating a KSADS interview or neurocognitive testing to check for loss of diagnosis or measurable change rarely makes sense in practice. Measures of mood severity are worth repeating if they are brief enough to be well-tolerated, yet still are valid.

  Assessment during treatment has many parallels to teaching. A "final exam" outcome evaluation should be comprehensive, covering not just key themes but also related material. A fairly intensive "midterm" assessment could evaluate progress and guide adjustments in the next phases. Brief and frequent evaluations - such as quizzes, homework assignments, and diaries - now have direct analogs in the mood assessment portfolio. Intensive interviews about the severity of mood, such as the CDRS-R (76) or KSADS Mania Rating Scale (62), provide valuable information about one aspect of functioning, but they require a considerable amount of time to complete and "grade" – similar to essays on a midterm or final. Their time demands prohibit frequent use in practice, and their narrow focus means that a comprehensive picture should augment via checklists (the "multiple choice" analog of the mood battery). Assessment best supports treatment by blending brief "process" measures (Step 8) with more intensive strategies quantifying severity or functioning at the beginning, middle, and end of treatment. Chronologically, these assessment tactics can weave together, just as Steps 8 and 9 are interdigitated here.

  Some behavior rating scales offer "good enough" validity in terms of sensitivity to severity and to treatment effects that they can be used instead of repeated semi-structured interviews in clinical practice (77, 78). Again, price and speed break ties between otherwise equally valid tools. The few studies investigating comparative treatment sensitivity find similar effect size estimates whether using parent-reported checklists versus semi-structured parent and youth interviews (78, 79). Parent checklists show stronger correlations than youth checklists do with criteria such as interview severity ratings or treatment effects. All things being equal, longer instruments will be more reliable, and reliability sets the upper limit on validity (80). Despite this psychometric principle, shorter versions of parent checklists are equally as sensitive to treatment effects as the full length versions, because they dropped weak items less specific to PBD. Thus, practitioners can use shorter versions as a baseline, midterm, and "final exam" and sacrifice little compared to doing a longer interview, and nothing compared to using a longer checklist.

  Practitioners should always try to involve the parent when working with mood disorder. This is standard practice working with young children, but becomes more variable when working with adolescents. Parents are helpful in identifying presence of mood disorder; they are sensitive to treatment effects even in blinded studies, and they are pivotal for retention in outpatient services. Most often the parent initiates the referral and provides transportation to sessions (81). If the parent does not feel heard, services often terminate prematurely.

  Despite the advantages of parent report for diagnosing PBD, there still is a major role for cross-informant perspectives in the context of treatment planning (74). Youth report, although second best to caregiver report at identifying mania (40, 47), is crucial for evaluating depression. Youth report also reveals the degree of insight and motivation for treatment (81). When parents report concerns that the youth denies or minimizes, then the practitioner and youth likely have discordant views about medication or therapy techniques. Compliance will be poor as a result. Teacher report is

not as helpful as parent or youth report for identifying PBD (48, 57), but it can be valuable for understanding school functioning and interventions planning. When multiple informants agree about the presence of mood problems, then severity is clearly much worse (50, 82).

How often should mood ratings be repeated? Although clinical trials use them weekly (78), there are diminishing returns. Once or twice – baseline and "final exam"- are good standard practice, and a "midterm" evaluation to make sure that there has been symptom reduction after an adequate trial of the therapy could be worthwhile. Assessments may detect early response and guide more rapid treatment adjustments (83), suggesting extra panels of assessment at the beginning of treatment or following major changes in regimen.

Clinicians often assess mood and energy informally during therapy or medication checks. Two simple modifications transform this into a powerful assessment strategy: (a) use a consistent scale, and (b) write it down (84). The choice of scale probably does not matter: A scale from 1 to 7, such as the Clinical Global Impressions scale uses (85), is simple enough that even children understand it, and more complex scales do not gain any sensitivity with their exaggerated appearance of precision. Writing down ratings over time exposes trends. A good minimum standard would be to incorporate a "mood and energy checkup" into the treatment note, so that each visit documents whether there has been a change in energy or mood during the intervening period (47).

Life charting is especially valuable for tracking treatment response, treatment emergent side effects, and triggering events linked to mood exacerbations (86). Life charts map naturally onto the "three column charts" and "five column charts" of classic cognitive behavioral therapy (87). Three column charts note times of intense emotional reaction, along with the triggering event and the attached cognition or interpretation. Life charts already record the emotion and the trigger, providing excellent source material for practicing the therapy skills.

If there were significant comorbidities, or multiple impaired domains, then the "midterm" examination may include either a broad instrument or select scales to check that the other concerns are responding to treatment. Often good PBD treatment reduces anxiety or attention problems, but sometimes these problems persist even when mood symptoms subside (59). The frequent persistence is a line of evidence suggesting that PBD and ADHD may be a true comorbidity (88). Stimulant medication can be well tolerated, particularly when mood stabilizers are already in place (59, 89). Assessment identifies if there are lingering symptoms that merit adjunctive treatment, and then monitors for treatment emergent changes in mood or energy. Other components for the mid-term exam may include assessment of family functioning, of substance use, or any other factors that might moderate treatment effects.

Measures of effect size are group statistics, not directly applicable to individual cases. There are several research definitions of treatment response and "clinically significant change" (90). Although these apply directly to individual cases, they have not become popular with practitioners. Practitioners tend to like people and not numbers; clinical significance definitions often set a high bar that is not achieved by many cases; and there is a fear of being evaluated by third party (or by consumers). However, without assessment we cannot learn from our mistakes (18); otherwise, in mental health, greater experience is not associated with better outcomes.

- **Step 10.** Maintenance monitoring. Once acute treatment concludes successfully, the last role for assessment is to provide early warning of potential relapse. Life charting and online methods can note changes in sleep or energy that might signal an incipient mood episode (28). Because these methods are novel, there are not yet clear evidence-based practice standards. The proactive practitioner develops a plan for self-monitoring, including identifying key triggers and warning signs that the person's mood may be "roughening" or destabilizing.

BENEFITS OF AN EVIDENCE BASED ASSESSMENT APPROACH TO PEDIATRIC BIPOLAR DISORDER

The assessment algorithm described here has several strengths. First, the approach is more accurate than unaided clinical decision making for PBD (21), replicating a well-established finding in literally hundreds of studies across clinical professions (18, 91). Second, information is used more consistently and efficiently. When clinicians read a vignette, hear about a family history of bipolarity, or see an elevated test score, they tend to attach different weights and meanings. An EBM approach weights the findings based on empirical validity. The more that clinicians use the EBM approach, the more consistent their interpretations of the same information will be, and the less contradictory opinions will result. Using the nomogram (a type of chart that yields Bayesian probabilities without

requiring any computation; see Appendix I) or other rapid interpretive approaches have shown dramatic increases in the precision of risk estimates for PBD (21). Third, the EBA approach eliminates a tendency to over-estimate the probability of PBD. The cognitive decision-making literature shows that humans focus more on cues of risk, and intuitively overestimate the probability of negative outcomes – erring on the side of caution for evolutionarily adaptive reasons (92). This evolved bias can lead to clinicians over-estimating the probability of negative outcomes such as PBD or suicide risk. EBM interprets objective inputs as objectively as possible, eliminating the potential for cognitive biases to distort the interpretation. Fourth, the algorithm is highly feasible. Thinking about the base rate, knowing the weight to assign to different risk factors, and knowing the information value linked to scores on broad-band assessment tools all add little or no time or cost to the intake process (see Table 1). It is a way of "working smarter," integrating these pieces of information into an "index of suspicion" that then guides the next clinical action. Adding a mania-specific measure also is highly feasible: The three measures with the strongest evidence base are also three of the shortest, and all are currently in the public domain. Thus adding a mania measure costs nothing, takes minimal time, and a one page version appears as helpful as longer or commercially distributed versions.

WHAT ARE FUTURE DIRECTIONS FOR EVIDENCE-BASED ASSESSMENT OF PEDIATRIC BIPOLAR DISORDER?

There has been remarkable progress in the assessment of PBD over the last two decades. There remains much to do before assessment "matures" to fully realize its potential for the identification and management of PBD. One frontier involves instrument translation and validation in languages other than English, and developing low cost tools that require minimal infrastructure, such as SMS text message mood charts. The next decade will bring decreases in the cost of technologies such as imaging and gene testing or proteomics, as well as advances in the delivery of computer-based performance measures.

None of these technologies or refinements will replace the clinician. A skilled professional remains essential to frame the questions of assessment, organize the tools, integrate the information, and interpret the data in a way that conveys meaning and motivation to the patient. Advances in assessment require new skills from the practitioner, the most central of which is the ability to balance and shift between technical and quantitative aspects of testing (27) and humanistic, qualitative aspects of interpretation (75). The methods presented here push for more systematic evaluation, incorporating validated tools, and shifting to a Bayesian framework for thinking about probability, risks, and benefits. To deliver any benefit, though, practitioners must develop competence and comfort with the concepts, so that they can explain findings in clear terms to a lay audience, and help patients to see how accurate assessment gives them power over their mood to change their lives for the better.

References

1. Carlson GA, Dyson M. Diagnostic implications of informant disagreement about rage outbursts: Bipolar disorder or another condition? Isr J Psychiatry Rel Sci 2012; 49: 44-51.
2. Dickstein DP, Leibenluft E. Beyond dogma: From diagnostic controversies to data about pediatric bipolar disorder and children with chronic irritability and mood dysregulation. Isr J Psychiatry Rel Sci 2012; 49: 52-61.
3. Goldstein B, Birmaher B. Prevalence, clinical presentation and differential diagnosis of pediatric bipolar disorder. Isr J Psychiatry Rel Sc 2012; 49: 3-14.
4. Youngstrom EA, Birmaher B, Findling RL. Pediatric bipolar disorder: Validity, phenomenology, and recommendations for diagnosis. Bipolar Disord 2008;10:194-214.
5. Van Meter A, Moreira AL, Youngstrom EA. Meta-analysis of epidemiological studies of pediatric bipolar disorder. J Clin Psychiatry 2011;72:1250-1256.
6. Kozloff N, Cheung AH, Schaffer A, Cairney J, Dewa CS, Veldhuizen S, et al. Bipolar disorder among adolescents and young adults: Results from an epidemiological sample. J Affect Disord 2010;125:350-354.
7. Moreno C, Laje G, Blanco C, Jiang H, Schmidt AB, Olfson M. National trends in the outpatient diagnosis and treatment of bipolar disorder in youth. Arch Gen Psychiatry 2007;64:1032-1039.
8. Lewinsohn PM, Seeley JR, Buckley ME, Klein DN. Bipolar disorder in adolescence and young adulthood. Child Adolesc Psychiatr Clin N Am 2002;11:461-476.
9. Merikangas KR, He JP, Burstein M, Swendsen J, Avenevoli S, Case B, et al. Service utilization for lifetime mental disorders in U.S. adolescents: Results of the National Comorbidity Survey-Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2011;50:32-45.
10. Birmaher B, Axelson D, Goldstein B, Strober M, Gill MK, Hunt J, et al. Four-year longitudinal course of children and adolescents with bipolar spectrum disorders: The Course and Outcome of Bipolar Youth (COBY) Study. Am J Psychiatry 2009;166:795-804.
11. Findling RL, Youngstrom EA, McNamara NK, Stansbrey RJ, Demeter CA, Bedoya D, et al. Early symptoms of mania and the role of parental risk. Bipolar Disord 2005;7:623-634.
12. Findling RL, Youngstrom EA, Fristad MA, Birmaher B, Kowatch RA, Arnold LE, et al. Characteristics of children with elevated symptoms of mania: The Longitudinal Assessment of Manic Symptoms (LAMS) Study. J Clin Psychiatry 2010;71:1664-1672.
13. Wozniak J, Biederman J, Kiely K, Ablon JS, Faraone S, Mundy E, et al. Mania-like symptoms suggestive of childhood-onset bipolar disorder in clinically referred children. J Am Acad Child Adolesc Psychiatry 1995;34:867-876.
14. Geller B, Tillman R, Bolhofner K, Zimerman B. Child bipolar I disorder: Prospective continuity with adult bipolar I disorder; characteristics of second and third episodes; predictors of 8-year outcome. Arch Gen Psychiatry 2008;65:1125-1133.
15. Angst J, Gamma A, Benazzi F, Ajdacic V, Eich D, Rossler W. Toward a re-definition of subthreshold bipolarity: Epidemiology and proposed

criteria for bipolar-II, minor bipolar disorders and hypomania. J Affect Disord 2003;73:133-146.

16. Fiedorowicz JG, Endicott J, Leon AC, Solomon DA, Keller MB, Coryell WH. Subthreshold hypomanic symptoms in progression from unipolar major depression to bipolar disorder. Am J Psychiatry 2011;168:40-48.
17. DelBello MP, Lopez-Larson MP, Soutullo CA, Strakowski SM. Effects of race on psychiatric diagnosis of hospitalized adolescents: A retrospective chart review. J Child Adolesc Psychopharmacol 2001;11:95-103.
18. Garb HN. Studying the clinician: Judgment research and psychological assessment. Washington, DC: American Psychological Association, 1998.
19. Biederman J, Klein RG, Pine DS, Klein DF. Resolved: Mania is mistaken for ADHD in prepubertal children. J Am Acad Child Adolesc Psychiatry 1998;37:1091-1093.
20. Klein RG, Pine DS, Klein DF. Resolved: Mania is mistaken for ADHD in prepubertal children. J Am Acad Child Adolesc Psychiatry 1998;37:1093-1096.
21. Jenkins MM, Youngstrom EA, Washburn JJ, Youngstrom JK. Evidence-based strategies improve assessment of pediatric bipolar disorder by community practitioners. Prof Psychol: Res Pract 2011;42:121-129.
22. Dubicka B, Carlson GA, Vail A, Harrington R. Prepubertal mania: Diagnostic differences between US and UK clinicians. Eur Child Adolesc Psychiatry 2008;17:153-161.
23. Rettew DC, Lynch AD, Achenbach TM, Dumenci L, Ivanova MY. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. Int J Methods Psychiatr Res 2009;18:169-184.
24. Marchand WR, Wirth L, Simon C. Delayed diagnosis of pediatric bipolar disorder in a community mental health setting. J Psych Prac 2006;12:128-133.
25. Lish JD, Dime-Meenan S, Whybrow PC, Price RA, Hirschfeld RM. The National Depressive and Manic-Depressive Association (DMDA) survey of bipolar members. J Affect Disord 1994;31:281-294.
26. Hirschfeld RM, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: How far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. J Clin Psychiatry 2003;64:161-174.
27. Straus SE, Glasziou P, Richardson WS, Haynes RB. Evidence-based medicine: How to practice and teach EBM. 4th ed. New York: Churchill Livingstone, 2011.
28. Sachs GS. Strategies for improving treatment of bipolar disorder: Integration of measurement and management. Acta Psychiatr Scand 2004:7-17.
29. Phelps J, Angst J, Katzow J, Sadler J. Validity and utility of bipolar spectrum models. Bipolar Disord 2008;10:179-193.
30. Kraemer HC. Evaluating medical tests: Objective and quantitative guidelines. Newbury Park, Cal.: Sage, 1992.
31. Hodgins S, Faucher B, Zarac A, Ellenbogen M. Children of parents with bipolar disorder. A population at high risk for major affective disorders. Child Adolesc Psychiatr Clin N Am 2002;11:533-553.
32. Algorta GP, Youngstrom EA, Phelps J, Jenkins MM, Youngstrom JK, Findling RL. An inexpensive family index of risk for mood issues improves identification of pediatric Bipolar Disorder Psychol Assess, Under review.
33. Youngstrom EA, Findling RL, Youngstrom JK, Calabrese JR. Toward an evidence-based assessment of pediatric bipolar disorder. J Clin Child Adolesc Psychol 2005;34:433-448.
34. Tillman R, Geller B, Klages T, Corrigan M, Bolhofner K, Zimerman B. Psychotic phenomena in 257 young children and adolescents with bipolar I disorder: Delusions and hallucinations (benign and pathological). Bipolar Disord 2008;10:45-55.
35. Murray G, Harvey A. Circadian rhythms and sleep in bipolar disorder. Bipolar Disord 2010;12:459-472.
36. Kowatch RA, Youngstrom EA, Danielyan A, Findling RL. Review and meta-analysis of the phenomenology and clinical characteristics of mania in children and adolescents. Bipolar Disord 2005;7:483-496.
37. Achenbach TM, Rescorla LA. Manual for the ASEBA School-Age Forms & Profiles. Burlington, Vermont: University of Vermont, 2001.
38. Reynolds CR, Kamphaus R. BASC-2 Behavior Assessment System for Children. Circle Pines, Minn.: American Guidance Service, 2004.
39. Mick E, Biederman J, Pandina G, Faraone SV. A preliminary meta-analysis of the child behavior checklist in pediatric bipolar disorder. Biol Psychiatry 2003;53:1021-1027.
40. Youngstrom EA, Findling RL, Calabrese JR, Gracious BL, Demeter C, DelPorto Bedoya D, et al. 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 2004;43:847-858.
41. Youngstrom EA, Meyers OI, Demeter C, Kogos Youngstrom J, Morello L, Piiparinen R, et al. Comparing diagnostic checklists for pediatric bipolar disorder in academic and community mental health settings. Bipolar Disord 2005;7:507-517.
42. Kahana SY, Youngstrom EA, Findling RL, Calabrese JR. 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 2003;13:471-488.
43. Diler RS, Birmaher B, Axelson D, Goldstein B, Gill M, Strober M, et al. The Child Behavior Checklist (CBCL) and the CBCL-bipolar phenotype are not useful in diagnosing pediatric bipolar disorder. J Child Adolesc Psychopharmacol 2009;19:23-30.
44. Wagner KD, Hirschfeld R, Findling RL, Emslie GJ, Gracious B, Reed M. Validation of the Mood Disorder Questionnaire for Bipolar Disorders in Adolescents. J Clin Psychiatry 2006;67:827-830.
45. Henry DB, Pavuluri MN, Youngstrom E, Birmaher B. Accuracy of brief and full forms of the Child Mania Rating Scale. J Clin Psychol 2008;64:368-381.
46. Youngstrom EA, Findling RL, Danielson CK, Calabrese JR. Discriminative validity of parent report of hypomanic and depressive symptoms on the General Behavior Inventory. Psychol Assess 2001;13:267-276.
47. Youngstrom EA, Freeman AJ, Jenkins MM. The assessment of children and adolescents with bipolar disorder. Child Adolesc Psychiatr Clin N Am 2009;18:353-390.
48. Hazell PL, Lewin TJ, Carr VJ. Confirmation that Child Behavior Checklist clinical scales discriminate juvenile mania from attention deficit hyperactivity disorder. J Paediatr Child Health 1999;35:199-203.
49. Geller B, Warner K, Williams M, Zimerman B. Prepubertal and young adolescent bipolarity versus ADHD: Assessment and validity using the WASH-U-KSADS, CBCL and TRF. J Affect Disord 1998;51:93-100.
50. Carlson GA, Youngstrom EA. Clinical implications of pervasive manic symptoms in children. Biol Psychiatry 2003;53:1050-1058.
51. Thuppal M, Carlson GA, Sprafkin J, Gadow KD. Correspondence between adolescent report, parent report, and teacher report of manic symptoms. J Child Adolesc Psychopharmacol 2002;12:27-35.
52. Youngstrom EA, Meyers O, Youngstrom JK, Calabrese JR, Findling RL. Diagnostic and measurement issues in the assessment of pediatric bipolar disorder: Implications for understanding mood disorder across the life cycle. Dev Psychopathol 2006;18:989-1021.
53. Achenbach TM, McConaughy SH, Howell CT. Child/Adolescent behavioral and emotional problems: Implication of cross-informant correlations for situational specificity. Psychol Bull 1987;101:213-232.
54. Loeber R, Green SM, Lahey BB. Mental health professionals' perception of the utility of children, mothers, and teachers as informants on childhood psychopathology. J Clin Child Psychol 1990;19:136-143.
55. Pini S, Dell'Osso L, Amador XF. Insight into illness in schizophrenia, schizoaffective disorder, and mood disorders with psychotic features. Am J Psychiatry 2001;158:122-125.
56. Dell'Osso L, Pini S, Cassano GB, Mastrocinque C, Seckinger RA, Saettoni M, et al. Insight into illness in patients with mania, mixed

mania, bipolar depression and major depression with psychotic features. Bipolar Disord 2002;4:315-322.

57. Youngstrom EA, Joseph MF, Greene J. Comparing the psychometric properties of multiple teacher report instruments as predictors of bipolar disorder in children and adolescents. J Clin Psychol 2008;64:382-401.
58. Youngstrom EA, Youngstrom JK, Freeman AJ, De Los Reyes A, Feeny NC, Findling RL. Informants are not all equal: Predictors and correlates of clinician judgments about caregiver and youth credibility. J Child Adolesc Psychopharmacol 2011;21:407-415.
59. McClellan J, Kowatch R, Findling RL. Practice parameter for the assessment and treatment of children and adolescents with bipolar disorder. J Am Acad Child Adolesc Psychiatry 2007;46:107-125.
60. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59:22-33.
61. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 1997;36:980-988.
62. Axelson DA, Birmaher BJ, Brent D, Wassick S, Hoover C, Bridge J, et al. A preliminary study of the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children mania rating scale for children and adolescents. J Child Adolesc Psychopharmacol 2003;13:463-470.
63. Suppiger A, In-Albon T, Hendriksen S, Hermann E, Margraf J, Schneider S. Acceptance of structured diagnostic interviews for mental disorders in clinical practice and research settings. Behav Ther 2009;40:272-279.
64. Jensen AL, Weisz JR. Assessing match and mismatch between practitioner-generated and standardized interview-generated diagnoses for clinic-referred children and adolescents. J Consult Clin Psychol 2002;70:158-168.
65. Pogge DL, Wayland-Smith D, Zaccario M, Borgaro S, Stokes J, Harvey PD. Diagnosis of manic episodes in adolescent inpatients: Structured diagnostic procedures compared to clinical chart diagnoses. Psychiatry Res 2001;101:47-54.
66. Mackin P, Targum SD, Kalali A, Rom D, Young AH. Culture and assessment of manic symptoms. Br J Psychiatry 2006;189:379-380.
67. Neighbors HW, Trierweiler SJ, Munday C, Thompson EE, Jackson JS, Binion VJ, et al. Psychiatric diagnosis of African Americans: Diagnostic divergence in clinician-structured and semistructured interviewing conditions. J Nat Med Assoc 1999;91:601-612.
68. Denicoff KD, Smith-Jackson EE, Disney ER, Suddath RL, Leverich GS, Post RM. Preliminary evidence of the reliability and validity of the prospective life-chart methodology (LCM-p). J Psychiatr Res 1997;31:593-603.
69. Youngstrom EA, Meyers OI, Youngstrom JK, Calabrese JR, Findling RL. Comparing the effects of sampling designs on the diagnostic accuracy of eight promising screening algorithms for pediatric bipolar disorder. Biol Psychiatry 2006;60:1013-1019.
70. Freeman AJ, Youngstrom EA, Michalak E, Siegel R, Meyers OI, Findling RL. Quality of life in pediatric bipolar disorder. Pediatrics 2009;123:e446-452.
71. Barnett JH, Huang J, Perlis RH, Young MM, Rosenbaum JF, Nierenberg AA, et al. Personality and bipolar disorder: Dissecting state and trait associations between mood and personality. Psychol Med 2011;41:1593-1604.
72. Ravens-Sieberer U, Bullinger M. Assessing health-related quality of life in chronically ill children with the German KINDL: First psychometric and content analytic results. Qual Life Res 1998;7:399-407.
73. Meyer RE, Salzman C, Youngstrom EA, Clayton PJ, Goodwin FK, Mann JJ, et al. Suicidality and risk of suicide - definition, drug safety concerns, and a necessary target for drug development: A brief report. J Clin Psychiatry 2010;71:1040-1046.
74. Youngstrom EA. Evidence-based strategies for the assessment of developmental psychopathology: Measuring prediction, prescription, and processs. In: Miklowitz DJ, Craighead WE, Craighead L, editors. Developmental psychopathology. New York: Wiley, 2008: pp. 34-77.
75. Finn SE, Tonsager ME. Information-gathering and therapeutic models of assessment: Complementary paradigms. Psychol Assess 1997;9:374-385.
76. Poznanski EO, Miller E, Salguero C, Kelsh RC. Preliminary studies of the reliability and validity of the Children's Depression Rating Scale. J Am Acad Child Psychiatry 1984;23:191-197.
77. West AE, Celio CI, Henry DB, Pavuluri MN. Child Mania Rating Scale-Parent Version: A valid measure of symptom change due to pharmacotherapy. J Affect Disord 2011;128:112-119.
78. Findling RL, McNamara NK, Gracious BL, Youngstrom EA, Stansbrey RJ, Reed MD, et al. Combination lithium and divalproex in pediatric bipolarity. J Am Acad Child Adolesc Psychiatry 2003;42:895-901.
79. Findling RL, McNamara NK, Youngstrom EA, Stansbrey R, Gracious BL, Reed MD, et al. Double-blind 18-month trial of lithium versus divalproex maintenance treatment in pediatric bipolar disorder. J Am Acad Child Adolesc Psychiatry 2005;44:409-417.
80. Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and use. 2nd ed. New York: Oxford University Press, 1995.
81. Yeh M, Weisz J. Why are we here at the clinic? Parent-child (dis) agreement on referral problems at outpatient treatment entry. J Consult Clin Psychol 2001;69:1018-1025.
82. Youngstrom EA, Findling RL, Calabrese JR. Who are the comorbid adolescents? Agreement between psychiatric diagnosis, parent, teacher, and youth report. J Abnorm Child Psychol 2003;31:231-245.
83. Howard KI, Moras K, Brill PL, Martinovich Z, Lutz W. Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. Am Psychol 1996;51:1059-1064.
84. Meehl PE. Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota, 1954.
85. National Institute of Mental Health. Rating scales and assessment instruments for use in pediatric psychopharmacology research. Psychopharmacol Bull 1985;21:839-843.
86. Post RM, Leverich GS, Altshuler LL, Frye MA, Suppes T, Keck PE, et al. Differential clinical characteristics, medication usage, and treatment response of bipolar disorder in the US versus The Netherlands and Germany. Int Clin Psychopharmacol 2011;26:96-106.
87. Newman CF, Leahy RL, Beck AT, Reilly-Harrington NA, Gyulai L. Bipolar disorder: A cognitive therapy approach. Washington, DC: American Psychological Association, 2002.
88. Youngstrom EA, Arnold LE, Frazier TW. Bipolar and ADHD Comorbidity: Both artifact and outgrowth of shared mechanisms. Clin Psychol 2010; 17:350-359.
89. Scheffer RE, Kowatch RA, Carmody T, Rush AJ. Randomized, placebo-controlled trial of mixed amphetamine salts for symptoms of comorbid ADHD in pediatric bipolar disorder after mood stabilization with divalproex sodium. Am J Psychiatry 2005;162:58-64.
90. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991;59:12-19.
91. Grove WM, Zald DH, Lebow BS, Snitz BE, Nelson C. Clinical versus mechanical prediction: A meta-analysis. Psychol Assess 2000;12:19-30.
92. Gigerenzer G, Goldstein DG. Reasoning the fast and frugal way: Models of bounded rationality. Psychol Rev 1996;103:650-669.
93. Naylor MW, Anderson TR, Kruesi MJ, Stoewe M, editors. Pharmacoepidemiology of bipolar disorder in abused and neglected state wards. Poster presented at the National Meeting of the American Academy of Child and Adolescent Psychiatry, 2002, October, San Francisco.
94. Biederman J, Faraone S, Mick E, Wozniak J, Chen L, Ouellette C, et al. Attention-deficit hyperactivity disorder and juvenile mania: An overlooked comorbidity? J Am Acad Child Adolesc Psychiatry 1996;35:997-1008.

95. Teplin LA, Abram KM, McClelland GM, Dulcan MK, Mericle AA. Psychiatric disorders in youth in juvenile detention. Arch Gen Psychiatry 2002;59:1133-1143.

96. Pliszka SR, Sherman JO, Barrow MV, Irick S. Affective disorder in juvenile offenders: A preliminary study. Am J Psychiatry 2000;157:130-132.

97. Blader JC, Carlson GA. Increased rates of bipolar disorder diagnoses among U.S. child, adolescent, and adult inpatients, 1996-2004. Biol Psychiatry 2007;62:107-114.

98. Smoller JW, Finn CT. Family, twin, and adoption studies of bipolar disorder. Am J Med Genet C Semin Med Genet 2003;123C:48-58.

99. Kochman FJ, Hantouche E, Ferrari P, Lancrenon S, Bayart D, Akiskal HS. Cyclothymic temperament as a prospective predictor of bipolarity and suicidality in children and adolescents with major depressive disorder. J Affect Disord 2005;85:181-189.

100. Birmaher B, Ryan ND, Williamson DE, Brent DA, Kaufman J, Dahl RE, et al. Childhood and adolescent depression: a review of the past 10 years. Part I. J Am Acad Child Adolesc Psychiatry 1996;35:1427-1439.

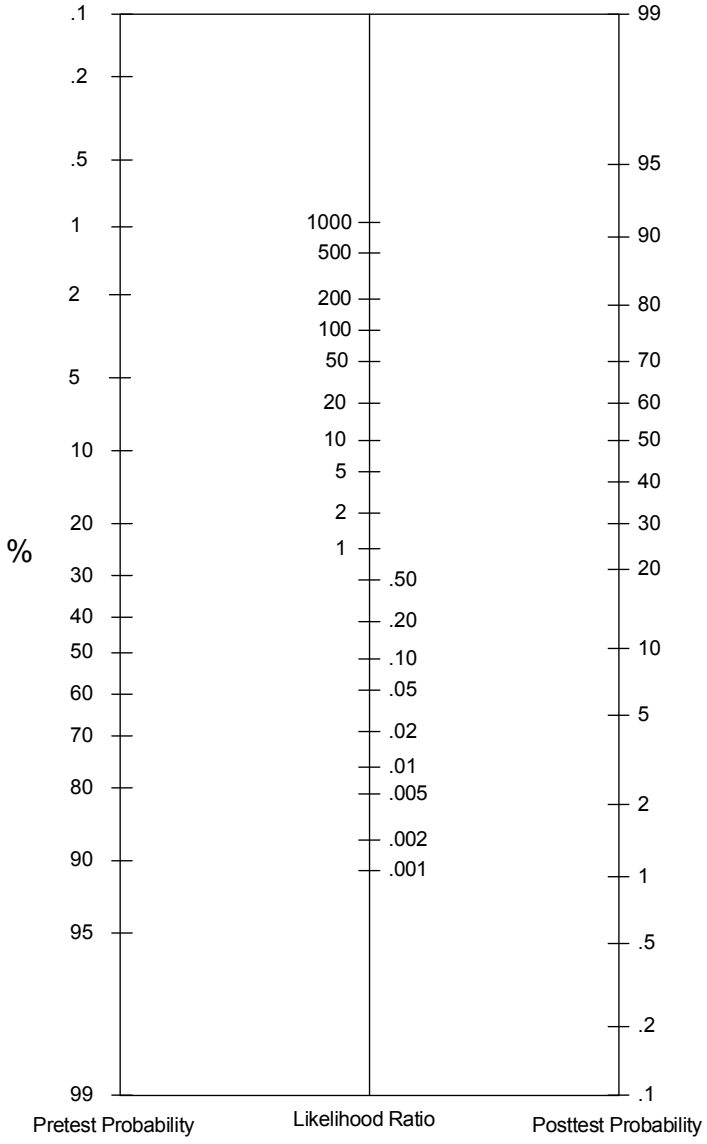
101. Joseph M, Youngstrom EA, Soares JC. Antidepressant-coincident mania in children and adolescents treated with selective serotonin reuptake inhibitors. Future Neurology 2009;4:87-102.

102. Achenbach TM. Manual for the Child Behavior Checklist/4-18 and 1991 profile. Burlington: University of Vermont, Department of Psychiatry, 1991.

103. Achenbach TM. Manual for the Youth Self Report form and 1991 profile. Burlington: University of Vermont, Department of Psychiatry, 1991.

104. Depue RA, Slater JF, Wolfstetter-Kausch H, Klein DN, Goplerud E, Farr DA. A behavioral paradigm for identifying persons at risk for bipolar depressive disorder: A conceptual framework and five validation studies. J Abnorm Psychol 1981;90:381-437.

Appendix 1. *Nomogram for combining probability with likelihood ratios*



Straus et al. (27) provide the rationale and examples of using the nomogram. Jenkins et al. (21) illustrate using a case with possible PBD.