Published in final edited form as:

Prof Psychol Res Pr. 2011 April ; 42(2): 121–129. doi:10.1037/a0022506.

Evidence-Based Strategies Improve Assessment of Pediatric Bipolar Disorder by Community Practitioners

Melissa M. Jenkins, MA¹, Eric A. Youngstrom, PhD¹, Jason J. Washburn, PhD^{2,3}, and Jennifer Kogos Youngstrom, PhD^{1,4}

¹ University of North Carolina at Chapel Hill

² Alexian Brothers Behavioral Health Hospital

³ Northwestern University Feinberg School of Medicine

⁴ Applewood Centers, Inc

Abstract

The misdiagnosis of pediatric bipolar disorder (PBD) has become a major public health concern. Would available evidence-based assessment (EBA) strategies help improve diagnostic accuracy, and are clinicians willing to consider these strategies in practice? The purpose of the present study was to document the extent to which using an EBA decision tool--a probability nomogram--improves the interpretation of family history and test data by clinicians and to examine the acceptability of the nomogram technique to clinicians. Over 600 clinicians across the US and Canada attending continuing education seminars were trained to use the nomogram. Participants estimated the probability that a youth in a clinical vignette had bipolar disorder, first using clinical judgment and then using the nomogram. Brief training of clinicians (less than 30 minutes) in using the nomogram for assessing PBD improved diagnostic accuracy, consistency, and agreement. The majority of clinicians endorsed using the nomogram in practice. EBA decision aids, such as the nomogram, may lead to a significant decrease in overdiagnosis and help clinicians detect true cases of PBD.

Keywords

Evidence-based assessment; pediatric bipolar disorder

How do clinicians respond when they are presented with identical case information about a child and then asked to determine the likelihood that the child has bipolar disorder? Take, for example, an 11-year-old African American male referred because of extreme aggression, distractibility, and motor agitation at school and who has a biological father diagnosed with bipolar I disorder and treated for several years with lithium and divalproex. What do clinicians estimate to be the likelihood of bipolar disorder for this child? Does the estimate change when clinicians are also informed that the mother reported a score of $T = 84$ on the

Correspondence concerning this article should be addressed to Eric A. Youngstrom, Professor of Psychology and Psychiatry, Associate Director, Center for Excellence in the Research & Treatment of Bipolar Disorder, University of North Carolina, CB #3270, Davie Hall, Chapel Hill, NC 27599-3270. eay@unc.edu.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/pro

Externalizing subscale of the Achenbach Child Behavior Checklist (CBCL) for this child (Achenbach & Rescorla, 2001)? Do clinicians interpret the same information similarly or differently? Is there a tendency to underestimate or overestimate the actual probability that the youth has bipolar disorder? Can evidence-based assessment (EBA) strategies help clinicians to interpret similar information similarly, improve the accuracy of the diagnosis of bipolar disorder, and help decrease underdiagnosis and overdiagnosis?

The State of Pediatric Bipolar Disorder

Providing a diagnosis of pediatric bipolar disorder (PBD) is controversial and challenging. PBD has received considerable attention in the research community and popular press (Kluger & Song, 2002; Papolos & Papolos, 1999). In the span of a decade, there was approximately a 10- to 40-fold increase in the diagnosis (Blader & Carlson, 2007) and treatment of bipolar in youths (Moreno et al., 2007). The rise in clinical diagnoses of PBD represents a major public health concern. Not only is bipolar the 6th leading cause of disability in adults (Murray, Lopez, & eds., 1996), it is associated with a 10 to 20 times increase in suicide risk compared to the general US population (Bostwick & Pankratz, 2000; Brodersen, Licht, Vestergaard, Olesen, & Mortensen, 2000; Guze & Robins, 1970; Harris & Barraclough, 1997; Sharma & Markar, 1994). Bipolar is also associated with substantial economic burden and medical conditions (Dunner, 2003; Kupfer, 2005; Murray et al., 1996; Stang et al., 2006).

The dramatic increase in clinical diagnosis of PBD raises the possibility that PBD may be overdiagnosed in many settings (Hirschfeld, Lewis, & Vornik, 2003; Kessler, Rubinow, Holmes, Abelson, & Zhao, 1997; Lish, Dime-Meenan, Whybrow, Price, & Hirschfeld, 1994). On the other hand, evidence shows that clinicians often take years to recognize bipolar disorder (Hirschfeld et al., 2003). For example, one study found that in over half of youth treated for bipolar disorder, at least five years elapsed from the onset of symptoms to a diagnosis (Marchand, Wirth, & Simon, 2006). Although there are concerns with overdiagnosis, clinicians may also miss true cases of PBD.

Misdiagnosis and delays in diagnosis carry serious consequences for patients, caregivers, and society. Youth with PBD who are misdiagnosed may receive ineffective or inappropriate treatment and follow a progressive and deteriorating course of bipolar illness (Geller, Tillman, Craney, & Bolhofner, 2004). Inappropriate pharmacologic treatment, such as antidepressants, is less effective than treatment with a mood stabilizer and can possibly worsen outcome (Altshuler et al., 1995; American Psychiatric Association, 2002; cf. Joseph, Youngstrom, & Soares, 2009; Hirschfeld et al., 2002; Sachs, Koslow, & Ghaemi, 2000). Conversely, diagnosing PBD when it is not present and thus unnecessarily starting pharmacological treatment for bipolar is dangerous, because medications used to treat the illness can have serious side effects (Wilens et al., 2003) including a potential increase in risk of suicidality (Goodwin et al., 2003).

Diagnostic Challenges

Correct diagnosis of PBD is crucial, but challenging (Bowring & Kovacs, 1992; Youngstrom, Findling, Youngstrom, & Calabrese, 2005). There have been important efforts to clarify the definitions of PBD (Leibenluft, Charney, Towbin, Bhangoo, & Pine, 2003; Youngstrom, 2009; Youngstrom, Birmaher, & Findling, 2008); however there is growing evidence that the use of different definitions of PBD matter in terms of course, neurocognitive functioning, and treatment response (Axelson et al., 2006; Birmaher et al., 2006; Leibenluft et al., 2003). Yet overlapping symptomatology makes it hard to tease out bipolar symptoms from symptoms of more prevalent diagnoses, such as Attention Deficit Hyperactivity Disorder (ADHD), unipolar depression, or conduct disorder (Bowring &

Kovacs, 1992; Kim & Miklowitz, 2002). Further, youths with PBD often meet criteria for other psychiatric disorders (Findling et al., 2001; Kowatch, Youngstrom, Danielyan, & Findling, 2005), making it difficult for clinicians to identify a “typical” PBD presentation in isolation. Complex presentation--coupled with the comparatively low prevalence of PBD--can lead clinicians either to focus on the comorbid condition and neglect PBD in treatment planning, or to misdiagnose cases with the cognate conditions as having PBD (Youngstrom et al., 2005). The varied presentations of PBD also threaten the reliability of diagnostic impressions and can make diagnostic decisions difficult. For example, classic bipolar I disorder can present as florid mania, severe depression, a mix of both, or as normal functioning, depending on the mood state. Evidence suggests that common clinical presentations often involve unstable mood and durations of hypomania or mania that are often shorter than the current DSM guidelines indicate, and this appears to be at least as frequently the case in youths as adults (Ghaemi et al., 2008; Youngstrom, Birmaher et al., 2008).

In addition to the challenges associated with the phenomenology of PBD, broader practice issues appear to complicate assessment. Research diagnostic instruments, such as structured and semi-structured interviews, can be impractical for use in clinical settings due to issues of training, burden, and reimbursement (Youngstrom, Meyers, Youngstrom, Calabrese, & Findling, 2006). Further, many common assessment methods are not evidence-based (as reviewed in Fletcher, Francis, Morris, & Lyon, 2005; Neisworth & Bagnato, 2004; Youngstrom et al., 2004; Youngstrom, Freeman, & Jenkins, 2009); and attempts to change clinician behavior often have not been successful (Galanter & Patel, 2005). Practitioners frequently rely on interviews and “unavoidably selective, reactive observations” for making important decisions (p. 202) (Peterson, 2004). EBA has lagged behind advances in other aspects of evidence-based practice (Mash & Hunsley, 2005), despite the development of techniques that could yield significant improvement (Jaeschke, Guyatt, & Sackett, 1994).

Clinical Judgment versus Actuarial Methods

Even when clinicians use valid and reliable behavioral checklists or tests, test interpretation is typically done using clinical judgment, which is prone to numerous errors (Croskerry, 2002; Elstein & Schwartz, 2002). Croskerry (2002) catalogs problematic decision-making shortcuts (heuristics) and biases, including descriptions and consequences for each. Diagnostic and treatment decisions are particularly vulnerable to faulty strategies and biases when clinicians rely solely on clinical judgment (Meehl, 1954). For example, the *availability heuristic*--the tendency to overestimate the frequency of an easily recalled event and underestimate the frequency of an ordinary or difficult to recall event--may lead clinicians to overdiagnose PBD due to the recent surge of media coverage on PBD (Galanter & Patel, 2005).

EBM advocates Bayesian approaches for assessing the probability that a patient has a particular disease, rather than relying solely on clinical judgment (Straus, Richardson, Glasziou, & Haynes, 2005). Bayesian approaches, such as the nomogram, use Baye's theorem to estimate the probability of a diagnosis based on test findings or clinical observations. Nomograms, which function like a probability slide rule, facilitate the estimation of probabilities without mathematical computation (see Figure 1). The nomogram is a simple, practical method for combining information about risk with the diagnostic likelihood ratios (DLRs) associated with test results or other clinical findings (Jaeschke et al., 1994). There are now multiple detailed examples of how this approach could be applied to psychiatric decision-making while incorporating information such as base rates, family history, behavior checklists, and performance based test results (Frazier & Youngstrom, 2006; Youngstrom & Duax, 2005; Youngstrom & Youngstrom, 2005).

Bayesian approaches have been available for centuries but have not gained popularity until recently, when EBM introduced actuarial methodologies in clinical settings. At present, the nomogram approach is prominently featured in clinical decision-making in EBM (Gray, 2004; Guyatt & Rennie, 2002; Straus et al., 2005); however, it is not taught in most psychiatric, psychological, or other mental health training programs. Thus, it seems crucial to compare how clinicians typically interpret test information with how they would perform using a nomogram.

The present study investigates the effectiveness of a specific EBA tool, the nomogram, in estimating the risk of PBD. Specific aims included: (1) to investigate whether using a nomogram improves the interpretation of family history and test data by clinicians; and (2) to examine the acceptability of the nomogram technique to clinicians.

Method

Participants

Participants were clinicians attending continuing education (CE) seminars about PBD in Toronto ($n=55$), Banff ($n=19$), Winnipeg ($n=48$), and London ($n=18$) in Canada; and Chapel Hill, North Carolina (3 sites) ($n=92$), Cleveland ($n=15$), Chicago (2 sites) ($n=120$), Ft. Wayne ($n=76$), Springfield ($n=53$), Virginia ($n=85$) and Orlando ($n=29$) in the USA, for a total $N=610$. Clinicians' professional titles ranged from licensed Master's level therapists to doctoral level psychologists, including newly licensed, mid-career, and late-career professionals.

Materials and Procedures

The primary outcome measure was estimates of the probability that a case vignette had PBD. The speaker presented a clinical vignette, and participants were asked to estimate the probability that the youth in the vignette had PBD based on DSM-IV criteria. The speaker then added a T -score on a widely used parent report norm-referenced behavioral checklist, with an associated diagnostic likelihood ratio (DLR) of 4.6. Participants accordingly revised their probability estimates. After estimating the likelihood of bipolar using clinical judgment, the speaker trained participants how to use the nomogram. After seeing a half-dozen worked examples, participants re-estimated the probability for the vignette based on the nomogram calculations. On average, the nomogram approach took participants less than 5 minutes. Data were collected anonymously. Because the data were anonymous and the vignette was a composite case, the institutional review board determined that these data did not constitute human subjects research under the purview of IRB oversight.

The clinical vignette included the CBCL. The test result was a highly elevated T Score on the CBCL Externalizing problems scale. Although more than a dozen instruments now have some research with regard to PBD (Youngstrom, 2007) and more specific measures for bipolar are available (see Youngstrom et al., 2005), we chose the CBCL because it is one of the most widely used and extensively researched instruments (Mick, Biederman, Pandina, & Faraone, 2003) and youth with PBD show significant elevations on the CBCL externalizing problems score (Mick et al., 2003; Youngstrom & Youngstrom, 2005).

Figure 2 illustrates the use of the nomogram with information in the clinical vignette; steps 1–6 show how participants were trained to use the nomogram. During the CE seminar, participants learned how to determine an appropriate starting base rate given the clinical setting, and how to translate family history and information from the CBCL into DLRs by referring to the literature. After synthesizing this information using clinical judgment, participants used the nomogram. Note that study participants used nomograms resembling Figure 1--without labels for settings (e.g., "outpatient mental health") and without the

translation of information into DLRs at the bottom of the middle columns of the nomogram (e.g., “1st degree = 5”)--and not Figure 2.

Statistical Analyses

Repeated measures ANOVA compared diagnostic impressions before and after training. Between groups ANOVA tested for differences across sites, and Levene's test of homogeneity compared the variances. Cohen's *d* measured effect size.

Results

Initial clinical judgment risk estimates of having PBD ranged from 0% to 100% ($M = 41\%$, $\pm 22\%$). Adding a test did not significantly narrow the range of opinion, which extended from 2% to 100% ($M = 60\%$, $\pm 21\%$). Average clinical estimates--with or without the test--were significantly higher than the Bayes' estimate of actual risk, both $p < .0005$.

Using the nomogram significantly reduced the mean and *SDs* of participants' risk estimates, all p values $< .0005$. The average risk estimate based on family history was 24% ($SD = 7\%$) when using the nomogram, repeated measures $F(1,596) = 343.87$ compared to clinical judgment; and 54% risk ($SD = 10\%$) after adding the test result, repeated measures $F(1,587) = 43.54$ compared to clinical judgment. The average estimate derived from base rate and family history information was no longer significantly different than the Bayes' estimate. The number of estimates within 5% of the actual Bayesian posterior probability rose from 19% to 88%.

Fewer than 5% of participants reported prior exposure to the nomogram or diagnostic likelihood ratios. Self assessment of knowledge about the new method for interpreting test results rose from 1.6 to 4.0 (1 = poor, 5 = excellent), with an effect size of $d = 3.2$ (where 0.8 is considered a “large” effect), $p < .0005$. After seeing feedback about the change in performance, 89% reported that they would consider using the nomogram in their own practice.

Average risk estimates also differed significantly between CE sites when participants relied solely on clinical judgment, $F(13,593) = 6.08$, $p < .0005$, and after adding a test, $F(13,594) = 10.27$, $p < .0005$. When participants used the nomogram to calculate the risk of having PBD (incorporating base rates and family history), CE sites' averages did not significantly differ, $F(13,586) = 1.67$, $p > .05$. Whereas 18.4% of the variance in clinical judgment ratings was attributable to between-site differences, only 5.6% of variance in nomogram estimates differed between sites.

Figure 3 illustrates clinicians' diagnostic impressions using clinical judgment and the nomogram first, combining base rate and family history and, second, incorporating test information. In contrast to the wide range of estimated probabilities of a bipolar diagnosis when clinicians used clinical judgment, the nomogram resulted in less variability in estimates around the “true” Bayesian probability. The gray band covers estimates lying within ± 5 percentage points around the Bayesian estimate (Sedlmeier & Gigerenzer, 2001).

Discussion

Study findings are consistent with qualms that clinicians may be prone to interpret identical information inconsistently (Garb, 1998) and to overdiagnose PBD when confronted with risk factors (Parens, Johnston, & Carlson, 2010). When participants estimated the risk of PBD using clinical judgment alone, interpretations ran the full gamut from 0% to 100% probability of PBD. Additional relevant assessment information from a norm-referenced test

did not significantly improve diagnostic accuracy or increase consensus. Taken together, these findings indicate that clinicians will often disagree in their diagnostic formulation of an individual case even when interpreting identical information (Dubicka, Carlson, Vail, & Harrington, 2008), and use of valid rating scales will not be sufficient by themselves to improve diagnostic accuracy.

The findings also indicate that the use of EBA strategies can lead to significant improvement in the interpretation of assessment information and reduce regional differences in decision-making. Community clinicians were able to use the nomogram to effectively combine multiple sources of information (e.g., base rate, familial risk, and test score) into probability estimates that were *consistent* (less spread in opinion), *unbiased* (neither systematically over- or under-estimating risk), and *efficient* (using a parsimonious amount of information to arrive at the estimate). A half-hour training produced large improvements in accuracy, as well as large effects in self-reported learning. Further, the majority of clinicians, almost all of whom were engaged primarily in clinical service delivery as opposed to research or administration, reported positive feedback about using the nomogram.

Bayesian approaches are a highly endorsed strategy for overcoming cognitive errors (Aegisdottir, et al., 2006; Arkes, 1981; Croskerry, 2003; Guyatt & Rennie, 2002; Straus, Richardson, Glasziou, & Haynes, 2005; Youngstrom, et al., 2009). In the case of bipolar disorder, these types of cognitive errors often result in misdiagnosis. Routine use of the nomogram for patients that present with symptoms of PBD, especially symptoms that overlap with other diagnoses (e.g., down mood, irritability, distractibility) can reduce the likelihood that bipolar disorder or a comorbid condition will be overlooked. Although the nomogram is not meant to be used in isolation to establish a diagnosis, it provides a relatively quick mechanism for ascertaining level of risk for PBD. Overall, the nomogram can be considered a recommended strategy for whenever a client is at elevated risk for PBD, such as when a client presents with manic symptoms, a family history of bipolar disorder, early onset depression or psychosis, or an elevated score on a relevant questionnaire (e.g., CBCL, PGBI) (Youngstrom, 2007).

Although the nomogram is one of many EBA strategies, it is unique for a number of reasons. First, costs are low. The nomogram itself is free, requires brief training, and takes relatively little time to execute in practice. Indeed, the clinicians included in this study came from diverse clinical backgrounds, suggesting that the nomogram can be implemented and positively received by a wide variety of clinicians. Second, the nomogram does not depend on having Internet access or using computer resources that may not always be available or practical in some settings. This is important, given that youths with mood disorders are more likely to be seen initially in primary care, family practice, or pediatric settings rather than specialty psychiatric clinics (e.g., Bickman & Noser, 2000).

Third, the nomogram is flexible and allows one to proceed with whatever pieces of information are available, unlike other clinical decision-making tools--such as logistic regression or decision trees--that one must abandon when any of the constituent predictor variables are missing. For example, although the study vignette in this study presented scores from the CBCL, the nomogram approach could integrate results from any test with diagnostic sensitivity and specificity estimates available. Recent reviews have collected evidence about available tests and calculated DLRs (Youngstrom, 2007; Youngstrom et al., 2009) to make it simpler for clinicians to use the nomogram and to choose instruments based on their validity for the question and patient at hand. Two of these instruments are free, in the public domain, and yield more decisive DLRs than the one used in the study vignette (Henry, Pavuluri, Youngstrom, & Birmaher, 2008; Youngstrom, Frazier, Demeter, Calabrese, & Findling, 2008).

Limitations

Like all heuristics, applying the nomogram method to PBD is not without limitations. First, the prevalence of PBD is a contentious topic, in part due to debate about the developmental appropriateness of current definitions of bipolar disorder (Leibenluft et al., 2003). Changing the definition of bipolar disorder, or including the spectrum of bipolar disorders (bipolar II, cyclothymic disorder, and bipolar not otherwise specified), alters the prevalence rate markedly (Youngstrom et al., 2009). However, improvements in definitions or prevalence estimates would also improve the quality of the data input into the nomogram, and thus refine the estimates. Changes in definitions would not prevent using this framework; instead, the nomogram approach ensures that available information is combined more accurately than would be accomplished with intuitive or unstructured approaches.

A second related limitation is that the use of local rates of PBD for initial base rates may be problematic if clinicians rely on idiosyncratic perceptions of PBD. Specifically, local base rates can under- or overestimate the actual base rate, depending on the extent to which clinicians employ conservative or “narrow” versus liberal or “broad” definitions. Being familiar with the prevalence of PBD at one’s clinic or similar settings and recognizing that base rates will likely change as the field progresses will counteract potential problems related to prevalence issues. EBP approaches also suggest performing “sensitivity analyses,” where the clinician examines the effect of using different base rates to see whether this changes clinical decision-making (Jaeschke et al., 1994; Straus et al., 2005). Recent reviews have gathered base rate estimates from multiple studies and clinics across a range of settings, providing reference values for comparison to local rates or for substitution into sensitivity analyses (Youngstrom et al., 2009).

A third potential technical limitation of the nomogram approach is that the pieces of information used to arrive at the probability estimate may sometimes lack independence. For example, using the nomogram to combine responses from the same parent on multiple different questionnaires is inappropriate because the questionnaire scores will be highly correlated and not contribute independent information. This is most likely to be a problem with multiple tests from the same informant. In contrast, the correlations between parent and youth or teacher report are sufficiently low that each perspective is likely to contribute distinct information. The nomogram approach also can synthesize information about clinical risk factors in combination with rating scale results. The nomogram approach has important advantages in terms of ease of use and flexibility in clinical settings compared to methods such as logistic regression that could also synthesize information adjusting for correlation among predictors.

Future Directions

Although initial response following training was positive, it is unclear whether front line clinicians are willing to use the nomogram in routine care over the long term. Little is known about potential barriers or more general policy considerations, including how to modify supervision, clinical training approaches, and reimbursement models to promote uptake of improved methods. More knowledge about these types of barriers can lead to more successful implementation efforts. Different mechanisms for “packaging” the nomogram method may also enhance the transportability of EBA tools into clinical arenas. For example, rather than using paper copies of the nomogram, software packages for computer or hand-held devices could be designed to automate the calculation process. Using technology to expedite the delivery of evidence-based services is a rapidly growing niche and can be an effective way for clinicians to adopt new assessment approaches. Despite the encouraging findings of the present study, more elaborate studies are needed, especially

those involving multiple case vignettes, to better understand the future role of Bayesian reasoning and the nomogram in clinical practice.

Decision support tools can improve screening and early detection of true cases of high risk illness, facilitating appropriate treatment while lessening both harm and economic burden (Esserman, Shieh, & Thompson, 2009). The nomogram may accomplish these specific goals in the assessment of PBD and other complicated diagnoses. As better tools become available, the nomogram provides a framework that will accept the upgrades and integrate the new test results or risk factors with other available information in a way that is flexible, individualized, and yet substantially more accurate than unaided decision-making using the same data. This framework also will be able to accommodate improvements in definitions and new evidence as it emerges about rates of disorders, risk factors, and empirically validated assessment procedures.

Although more research is recommended to further validate the use and attractiveness of the nomogram in real-world practice, findings from the present study suggest that actuarial methods can produce immediate improvement. For example, DLRs and prevalence rates from the current literature yield relatively conservative probabilities of a bipolar diagnosis even with high test scores and when family history is positive for bipolar (as was the case for the youth in the study vignette). Thus, the nomogram approach can decrease overdiagnosis and prematurely starting medication as well as anchor clinicians' judgments and reduce prematurely ruling out bipolar--helping to avoid two frequent and serious diagnostic errors in current practice.

Acknowledgments

Funding/Support: This work was supported in part by NIH R01 MH066647 (PI: E. Youngstrom) and the Jean Paul Ohadi Foundation (PI: J. Washburn).

References

- Achenbach, TM.; Rescorla, L. Manual for the ASEBA school-age forms & profiles. ASEBA; Burlington, VT: 2001.
- Aegisdottir S, White MJ, Spengler PM, Maugherman AS, Anderson LA, Cook RS, et al. The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*. 2006; 34(3):341.10.1177/0011000005285875
- Altshuler LL, Post RM, Leverich GS, Mikalauskas K, Rosoff A, Ackerman L. Antidepressant-induced mania and cycle acceleration: a controversy revisited. *The American Journal of Psychiatry*. 1995; 152(8):1130–1138. [PubMed: 7625459]
- American Psychiatric Association. Practice guideline for the treatment of patients with bipolar disorder (revision). *The American Journal of Psychiatry*. 2002; 159:1–50.
- Arkes HR. Impediments to accurate clinical judgment and possible ways to minimize their impact. *Journal of Consulting and Clinical Psychology*. 1981; 49(3):323–330.10.1037/0022-006X.49.3.323 [PubMed: 7276321]
- Axelson D, Birmaher B, Strober M, Gill MK, Valeri S, Chiappetta L, et al. Phenomenology of children and adolescents with bipolar spectrum disorders. *Archives of General Psychiatry*. 2006; 63(10): 1139–1148.10.1001/archpsyc.63.10.1139 [PubMed: 17015816]
- Bickman L, Noser K. Meeting the challenges in the delivery of child and adolescent mental health services in the next millenium: The continuous quality improvement approach. *Applied and Preventive Psychology*. 2000; 8(4)10.1016/S0962-1849(05)80039-3
- Birmaher B, Axelson D, Strober M, Gill MK, Valeri S, Chiappetta L, et al. Clinical course of children and adolescents with bipolar spectrum disorders. *Archives of General Psychiatry*. 2006; 63:175–183.10.1001/archpsyc.63.2.175 [PubMed: 16461861]

- Blader JC, Carlson GA. Increased rates of bipolar disorder diagnoses among U.S. child, adolescent, and adult inpatients, 1996–2004. *Biological Psychiatry*. 2007; 62(2):107–114.10.1016/j.biopsych.2006.11.006 [PubMed: 17306773]
- Bostwick JM, Pankratz VS. Affective disorders and suicide risk: a reexamination. *The American Journal of Psychiatry*. 2000; 157(12):1925–1932.10.1176/appi.ajp.157.12.1925 [PubMed: 11097952]
- Bowring MA, Kovacs M. Difficulties in diagnosing manic disorders among children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1992; 31(4):611–614.10.1097/00004583-199207000-00006 [PubMed: 1644722]
- Brodersen A, Licht RW, Vestergaard P, Olesen AV, Mortensen PB. Sixteen-year mortality in patients with affective disorder commenced on lithium. *The British Journal of Psychiatry*. 2000; 176:429.10.1192/bjp.176.5.429 [PubMed: 10912217]
- Croskerry P. Achieving quality in clinical decision making: cognitive strategies and detection of bias. *Academic Emergency Medicine*. 2002; 9(11):1184–1204.10.1111/j.1553-2712.2002.tb01574.x [PubMed: 12414468]
- Dubicka B, Carlson GA, Vail A, Harrington R. Prepubertal mania: diagnostic differences between US and UK clinicians. *European Child & Adolescent Psychiatry*. 2008; 17(3):153–161.10.1007/s00787-007-0649-5 [PubMed: 17876503]
- Dunner DL. Clinical consequences of under-recognized bipolar spectrum disorder. *Bipolar Disorders*. 2003; 5(6):456–463.10.1046/j.1399-5618.2003.00073.x [PubMed: 14636371]
- Elstein AS, Schwartz A. Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. *British Medical Journal*. 2002; 324(7339):729–732.10.1136/bmj.324.7339.729 [PubMed: 11909793]
- Esserman L, Shieh Y, Thompson I. Rethinking Screening for Breast Cancer and Prostate Cancer. *The Journal of the American Medical Association*. 2009; 302(15):1685.10.1001/jama.2009.1498
- Findling RL, Gracious BL, McNamara NK, Youngstrom EA, Demeter C, Calabrese JR. Rapid, continuous cycling and psychiatric co-morbidity in pediatric bipolar I disorder. *Bipolar Disorders*. 2001; 3:202–210.10.1034/j.1399-5618.2001.30405.x [PubMed: 11552959]
- Fletcher JM, Francis DJ, Morris RD, Lyon GR. Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child & Adolescent Psychology*. 2005; 34(3):506–522.10.1207/s15374424jccp3403_7 [PubMed: 16083392]
- Frazier TW, Youngstrom EA. Evidence Based Assessment of Attention-Deficit/Hyperactivity Disorder: Using multiple sources of information. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2006; 45:614–620.10.1097/01.chi.0000196597.09103.25 [PubMed: 16670656]
- Galanter CA, Patel VL. Medical decision making: a selective review for child psychiatrists and psychologists. *Journal of Child Psychology & Psychiatry*. 2005; 46(7):675–689.10.1111/j.1469-7610.2005.01452.x [PubMed: 15972065]
- Garb, HN. *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: American Psychological Association; 1998.
- Geller B, Tillman R, Craney JL, Bolhofner K. Four-year prospective outcome and natural history of mania in children with a prepubertal and early adolescent bipolar disorder phenotype. *Archives of General Psychiatry*. 2004; 61(5):459–467.10.1001/archpsyc.61.5.459 [PubMed: 15123490]
- Ghaemi SN, Bauer M, Cassidy F, Malhi GS, Mitchell P, Phelps J, et al. Diagnostic guidelines for bipolar disorder: a summary of the International Society for Bipolar Disorders Diagnostic Guidelines Task Force Report. *Bipolar Disorders*. 2008; 10(1 Pt 2):117–128.10.1111/j.1399-5618.2007.00556.x [PubMed: 18199230]
- Goodwin FK, Fireman B, Simon GE, Hunkeler EM, Lee J, Revicki D. Suicide Risk in Bipolar Disorder During Treatment With Lithium and Divalproex. *The Journal of the American Medical Association*. 2003; 290(11):1467–1473.10.1001/jama.290.11.1467
- Gray, GE. *Concise guide to evidence-based psychiatry*. American Psychiatric Pub; 2004.
- Guyatt, G.; Rennie, D. *Users' guides to the medical literature: essentials of evidence-based clinical practice*. Chicago: American Medical Association Press; 2002.

- Guze SB, Robins E. Suicide and primary affective disorders. *The British Journal of Psychiatry*. 1970; 117(539):437–438.10.1192/bjp.117.539.437 [PubMed: 5481206]
- Harris EC, Barraclough B. Suicide as an outcome for mental disorders. A meta-analysis. *The British Journal of Psychiatry*. 1997; 170(3):205.10.1192/bjp.170.3.205 [PubMed: 9229027]
- Henry DB, Pavuluri MN, Youngstrom EA, Birmaher B. Accuracy of brief and full forms of the child mania rating scale. *Journal of Clinical Psychology*. 2008; 64(4):368–381.10.1002/jclp.20464 [PubMed: 18302291]
- Hirschfeld, RM.; Bowden, CL.; Gitlin, MJ.; Keck, PE.; Suppes, T.; Thase, ME., et al. American Psychiatric Association practice guidelines for the treatment of psychiatric disorders: Compendium 2002. 2. Washington, DC US: American Psychiatric Association; 2002. Practice guideline for the treatment of patients with bipolar disorder; p. 547-634.
- 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. *The Journal of Clinical Psychiatry*. 2003; 64(2):161–174.10.4088/JCP.v64n0209 [PubMed: 12633125]
- Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature: III. How to use an article about a diagnostic test: A. Are the results of the study valid? *The Journal of the American Medical Association*. 1994; 271(5):389–391.10.1001/jama.1994.03510290071040
- Joseph MF, Youngstrom EA, Soares JC. Antidepressant-coincident mania in children and adolescents treated with selective serotonin reuptake inhibitors. *Future Neurology*. 2009; 4(1):87–102.10.2217/14796708.4.1.87 [PubMed: 19884978]
- Kessler, RC. Comorbidity of unipolar and bipolar depression with other psychiatric disorders in a general population survey. In: Tohen, M., editor. *Comorbidity in Affective Disorders*. New York: Marcel Dekker, Inc; 1999. p. 1-25.
- Kessler RC, Rubinow DR, Holmes C, Abelson JM, Zhao S. The epidemiology of DSM-III-R bipolar I disorder in a general population survey. *Psychological Medicine*. 1997; 27(5):1079–1089.10.1017/S0033291797005333 [PubMed: 9300513]
- Kim EY, Miklowitz DJ. Childhood mania, attention deficit hyperactivity disorder and conduct disorder: a critical review of diagnostic dilemmas. *Bipolar Disorders*. 2002; 4(4):215–225.10.1034/j.1399-5618.2002.01191.x [PubMed: 12190710]
- Kluger J, Song S. Young and bipolar. *Time*. 2002 August 19.:39–47. 51. [PubMed: 11951314]
- 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 Disorders*. 2005; 7:483–496.10.1111/j.1399-5618.2005.00261.x [PubMed: 16403174]
- Kupfer DJ. The increasing medical burden in bipolar disorder. *The Journal of the American Medical Association*. 2005; 293(20):2528–2530.10.1001/jama.293.20.2528
- Leibenluft E, Charney DS, Towbin KE, Bhangoo RK, Pine DS. Defining clinical phenotypes of juvenile mania. *The American Journal of Psychiatry*. 2003; 160(3):430.10.1176/appi.ajp.160.3.430 [PubMed: 12611821]
- Lish JD, Dime-Meenan S, Whybrow PC, Price RA, Hirschfeld RM. The National Depressive and Manic-depressive Association (DMDA) survey of bipolar members. *Journal of Affective Disorders*. 1994; 31(4):281–294.10.1016/0165-0327(94)90104-X [PubMed: 7989643]
- Marchand WR, Wirth L, Simon C. Delayed diagnosis of pediatric bipolar disorder in a community mental health setting. *Journal of Psychiatry Practice*. 2006; 12(2):128–133.10.1097/00131746-200603000-00011
- Mash EJ, Hunsley J. Evidence-Based Assessment of Child and Adolescent Disorders: Issues and Challenges. *Journal of Clinical Child & Adolescent Psychology*. 2005; 34(3):362–379.10.1207/s15374424jccp3403_1 [PubMed: 16026210]
- Meehl, PE. Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota Press; 1954.
- Mick E, Biederman J, Pandina G, Faraone SV. A preliminary meta-analysis of the child behavior checklist in pediatric bipolar disorder. *Biological Psychiatry*. 2003; 53(11):1021–1027.10.1016/S0006-3223(03)00234-8 [PubMed: 12788247]

- 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. *Archives of General Psychiatry*. 2007; 64(9): 1032–1039.10.1001/archpsyc.64.9.1032 [PubMed: 17768268]
- Murray, C.; Lopez, A., editors. *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Disease, Injuries, and Risk Factors in 1990 and Projected to 2020*. Cambridge, MA: Harvard University Press; 1996.
- Neisworth JT, Bagnato SJ. The Mismeasure of Young Children: The Authentic Assessment Alternative. *Infants and Young Children*. 2004; 17(3):198–212.10.1097/00001163-200407000-00002
- Papolos, DF.; Papolos, J. *The bipolar child: The definitive and reassuring guide to childhood's most misunderstood disorder*. New York: Broadway Books; 1999.
- Parens E, Johnston J, Carlson GA. Pediatric Mental Health Care Dysfunction Disorder? *New England Journal of Medicine*. 2010; 362(20):10.1056/NEJMp1003175
- Peterson DR. Science, Scientism, and Professional Responsibility. *Clinical Psychology: Science and Practice*. 2004; 11(2):196–210.10.1093/clipsy/bph072
- Sachs GS, Koslow CL, Ghaemi SN. The treatment of bipolar depression. *Bipolar Disorders*. 2000; 2(3p2):256–260.10.1034/j.1399-5618.2000.20306.x [PubMed: 11249803]
- Sedlmeier P, Gigerenzer G. Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*. 2001; 130(3):380–400.10.1037/0096-3445.130.3.380 [PubMed: 11561916]
- Sharma R, Markar HR. Mortality in affective disorder. *Journal of Affective Disorders*. 1994; 31(2):91–96.10.1016/0165-0327(94)90112-0 [PubMed: 8071480]
- Stang PE, Frank C, Kalsekar A, Yood MU, Wells K, Burch S. The clinical history and costs associated with delayed diagnosis of bipolar disorder. *Medscape General Medicine*. 2006; 8(2):18. [PubMed: 16926757]
- Straus, SE.; Richardson, WS.; Glasziou, P.; Haynes, RB. *Evidence-based medicine: How to practice and teach EBM*. 3. New York: Churchill Livingstone; 2005.
- Wilens TE, Biederman J, Kwon A, Chase R, Greenberg L, Mick E, et al. A systematic chart review of the nature of psychiatric adverse events in children and adolescents treated with selective serotonin reuptake inhibitors. *Journal of Child and Adolescent Psychopharmacology*. 2003; 13(2):143–152.10.1089/10445460322163862 [PubMed: 12886909]
- Youngstrom, EA. Pediatric Bipolar Disorder. In: Mash, EJ.; Barkley, RA., editors. *Assessment of Childhood Disorders*. New York: Guilford Press; 2007. p. 253-304.
- Youngstrom EA. Definitional issues in bipolar disorder across the life cycle. *Clinical Psychology Science Practice*. 2009; 16(2):140–160.10.1111/j.1468-2850.2009.01154.x
- Youngstrom EA, Birmaher B, Findling RL. Pediatric bipolar disorder: validity, phenomenology, and recommendations for diagnosis. *Bipolar Disorders*. 2008; 10:194–214.10.1111/j.1399-5618.2007.00563.x [PubMed: 18199237]
- Youngstrom EA, Duax J. Evidence based assessment of pediatric bipolar disorder, part 1: Base rate and family history. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2005; 44:712–717.10.1097/01.chi.0000162581.87710.bd [PubMed: 15968241]
- Youngstrom EA, Findling RL, Calabrese JR, Gracious BL, Demeter C, Bedoya DDP, et al. Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5 to 17 years. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2004; 43:847.10.1097/01.chi.0000125091.35109.1e [PubMed: 15213586]
- Youngstrom EA, Findling RL, Youngstrom JK, Calabrese JR. Toward an evidence-based assessment of pediatric bipolar disorder. *Journal of Clinical Child & Adolescent Psychology*. 2005; 34(3): 433–448.10.1207/s15374424jccp3403_4 [PubMed: 16026213]
- Youngstrom EA, Frazier TW, Demeter C, Calabrese JR, Findling RL. Developing a 10-item mania scale from the Parent General Behavior Inventory for children and adolescents. *The Journal of Clinical Psychiatry*. 2008; 69(5):831–839.10.4088/JCP.v69n0517 [PubMed: 18452343]
- Youngstrom EA, Freeman AJ, Jenkins MM. The assessment of children and adolescents with bipolar disorder. *Child & Adolescent Psychiatric Clinics of North America*. 2009; 18(2):353–390. viii–ix. 10.1016/j.chc.2008.12.002 [PubMed: 19264268]

- 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. *Development and Psychopathology*. 2006; 18(04):989–1021. 10.1017/S0954579406060494. [PubMed: 17064426]
- Youngstrom EA, Youngstrom JK. Evidence-based assessment of pediatric bipolar disorder, Part II: Incorporating information from behavior checklists. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2005; 44(8):823.10.1097/01.chi.0000164589.10200.a4 [PubMed: 16034285]

Biographies

ERIC A. YOUNGSTROM received his PhD from the University of Delaware. He is a professor of psychology and psychiatry at the University of North Carolina at Chapel Hill, where he is also the Acting Director of the Center for Excellence in Research and Treatment of Bipolar Disorder. His research interests include development and dissemination of evidence based assessment strategies, as well as understanding the phenomenology and course of bipolar disorder.

MELISSA M. JENKINS, MA, is a doctoral student at the University of North Carolina at Chapel Hill. She is a fourth year student in the clinical psychology program. Her areas of research and practice include clinical decision-making, community mental health, childhood mood disorders, and implementation and dissemination of evidence-based practices.

JASON J. WASHBURN received his PhD from DePaul University. He is Director of Education and Clinical Training for the Division of Psychology at the Feinberg School of Medicine at Northwestern University. He also is the director of the Center for Evidence-Based Practice (CEBP) at Alexian Brothers Behavioral Health Hospital, board-certified in clinical child and adolescent psychology (ABPP). His research interests include effectiveness research and the dissemination of evidence-based practices.

JENNIFER K. YOUNGSTROM received her PhD from the University of Delaware. She is the Director of Child and Family Services at the Psychology Department of the University of North Carolina at Chapel Hill. Her research interests include adapting and improving cognitive-behavioral interventions for work with diverse families, outcome evaluation, training and supervision.

Figure 1. Nomogram for Combining Probability with Diagnostic Likelihood Ratios

Note. A nomogram is a particularly helpful tool for quantifying the risk of bipolar disorder when certain warning signs are present (Youngstrom et al., 2009). Warning signs can include family history of bipolar disorder, a high score on a parent report questionnaire that is sensitive to manic symptoms, and/or a youth's clinical presentation of decreased need for sleep, elevated, expansive mood, and grandiosity, or possible psychotic features.

Figure 2. How to Use the Nomogram

Vignette: 7 year old African American male, referred because of extreme aggression and distractibility, and motor agitation at school. Biological father has been diagnosed with bipolar I disorder and treated for several years with lithium and divalproex. Mom completes Achenbach Child Behavior Checklist and he earns an Externalizing $T = 84$.

Note. Steps in using the nomogram: (1) Select appropriate pretest probability, typically the base rate of the disorder, absent any other information. Base rate information for pediatric bipolar disorder can be found in the literature (e.g., Youngstrom, et al., 2009, Table 2). (2) Find the Diagnostic Likelihood Ratio (DLR) associated with the risk factor or test result and plot it on the middle line. Bipolar disorder in a first degree relative has a DLR of 5.0. (3) Connect the dots and extend across the third line to estimate the posterior probability. (4) To add new information, repeat the process by using the posterior value from Step #3 as the new starting point. (5) Plot the DLR associated with the additional test result or risk factor on the middle line. (6) Connect the dots and extend across the third line to obtain the revised posterior probability. The order in which one enters the DLR information into the nomogram does not matter (e.g., one could first combine base rate and test score information and then family history information--the opposite order of what is illustrated above--and still arrive at the same posterior probability). One can also multiply the DLR's together (e.g., $5 * 4.6$ using the example above) instead of treating as separate pieces of information and plotting on two nomograms. Results are algebraically equivalent. For more information about the nomogram procedure, see Straus, Richardson, Glasziou, & Haynes (2005).

Figure 3. Comparison of Practitioner Estimates Using Clinical Judgment or the Nomogram. Gray bars indicate the range within $\pm 5\%$ of the Bayesian Estimates (24% and 54% for the two sets of clinical information).