Evidence-Based Assessment of Pediatric Bipolar Disorder, Part I: Base Rate and Family History

ERIC A. YOUNGSTROM, PH.D. AND JEANNE DUAX, B.A.

The objective of this column is to illustrate clinical decision making about the risk of pediatric bipolar disorder (PBD) using two key pieces of information: base rates and family history. A second commentary in next month's *Journal* details how to find and use behavior checklists as an additional source of information. The approach in both commentaries illustrates how to conduct focused database searches to address clinical questions and how to use so-called Bayesian methods to combine information from more than one source regarding the probability of a disorder. Combining clinical information provides a rational framework for deciding when to use additional and more expensive assessment methods (i.e., the test/no-test threshold) as well as when to initiate treatments specific to bipolar disorder (i.e., the Treatment Threshold [Guyatt and Rennie, 2002]). The following sections use a case example to model the process of gathering information and integrating it according to the recommendations of evidence-based practice (EBP). Both the challenges and the potential benefits of an EBP approach are magnified when considering a controversial and high-stakes diagnosis such as PBD. Some guidelines are developed that would be readily applicable to other diagnoses, and some specific recommendations about assessment instruments and strategies for bipolar disorder are provided.

Accepted February 11, 2005.

From the Departments of Psychology and Psychiatry, Case Western Reserve University/University Hospitals of Cleveland.

This research was supported by NIMH R01 MH-066647 as well as NIMH P20 MH 066054 and a Clinical Research Center Grant from the Stanley Medical Research Institute. Thanks also to John Hamilton, M.D., and Martha Hellander, J.D., for suggestions and comments.

Reprint requests to Dr. Eric A. Youngstrom, Department of Psychology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7123; e-mail to Eric.Youngstrom@case.edu.

DOI: 10.1097/01.chi.0000162581.87710.bd

CLINICAL VIGNETTE

A colleague works at a community mental health center in an urban setting. She receives a referral for a 9-year-old African-American boy (the specific details of the case have been changed so that the person described is not recognizable). The presenting problems were episodes of extreme aggression, along with difficulty concentrating and high levels of motor activity. The intake information suggests that both the frequency and intensity of the aggressive behavior seem unusual for a 9-year-old. For example, the child's mother reported that after chasing the family cat with scissors, he tried to stab her, and when she attempted to take the scissors away from him, he slammed his door so hard that the hinges tore loose. Your colleague is also concerned because there is a family history of bipolar disorder: The biological father has carried a diagnosis of bipolar I for several years and currently is stable on lithium monotherapy. The colleague is considering the possibility of a diagnosis of bipolar disorder in the child but is uncertain about how best to use the available information or proceed with the evaluation.

The colleague has good reason to be cautious. Bipolar disorder is probably rare in children before puberty, there is controversy about how to diagnose it, and there are few published clinical trials to guide treatment (Weckerly, 2002). There also are concerns that untreated bipolar disorder will follow a progressive and deteriorating course (Geller et al., 2004) and that use of stimulants or antidepressants might actually worsen the course of illness (Carlson, 2003). On the other hand, the compounds most likely to be helpful in treating bipolar disorder also have the potential for serious side effects, and thus they should not be prescribed to children unless one is fairly confident in the diagnosis and the potential for benefit (Weller et al., 2004).

HOW OFTEN DOES BIPOLAR DISORDER OCCUR IN ONE'S PRACTICE? ESTABLISHING THE BASE RATE

A crucial thing to know at the outset of the assessment process is how often different diagnoses occur in settings similar to where one works. Diagnosis in EBP is a *probability* of having the condition, a concept that is different from how diagnosis is thought about in usual care. The base rate of a condition provides an excellent starting probability for the EBP diagnosis of an individual patient. EBP also suggests that clinicians initially limit the diagnoses under consideration to (a) those that are most likely at a given setting (a *probabilistic* approach), (b) those that have the most serious consequences if left undiagnosed and untreated (a *prognostic* approach), and (c) those that are more responsive to treatment if offered (a *pragmatic* approach) (Guyatt and Rennie, 2002).

Ideally, we would know the frequency of various diagnoses in our specific practice. It may be possible to approximate these figures by reviewing the diagnoses assigned to one's patients over a period of time, or perhaps the agency uses a computer database that would allow reviewing diagnoses for all patients seen in a particular service. There are advantages and drawbacks to using local data, as detailed in Table 1. Clinical diagnoses need to be used with caution as a source of base rates: They often underestimate rates of comorbidity and can be affected by unique referral patterns or conceptualizations of a disorder. Local base rates need to be treated with extra caution when they diverge from published rates in comparable settings, unless there is a clear reason for the difference. These shortcomings are often magnified when considering PBD, where even experts vary in their description of mania and the relative emphasis placed on different signs and symptoms (Leibenluft et al., 2003). There also have been dramatic changes in the rate at which bipolar disorder is diagnosed over the past decade and in different regions of the country (see Youngstrom et al., in press a, for review).

Because of these limitations, it also would be valuable to have benchmarks indicating the frequency with which bipolar disorder is being diagnosed in different clinical settings. The best benchmarks for clinical use would be those that were developed and published recently. The benchmark also should derive from a sample that is similar to the target clinical setting in terms of demography, pattern of ascertainment, severity of impairment, and pattern of presenting diagnoses. Most of the base rate data on PBD have been based on epidemiological studies, which typically have enrolled mostly European-American participants (see Kessler et al., 2001, for review of issues in child studies) and not found sufficient cases with bipolar disorder to afford meaningful subgroup analyses by race, gender, ethnicity, or socioeconomic status. An additional concern with epidemiological studies is that they often rely on structured diagnostic interviews that have high reliability but may be less valid for detecting mania (e.g., Kessler et al., 1997), particularly if a collateral informant such as a parent is not routinely involved in the diagnostic interview (cf. Youngstrom et al., 2004). Table 2 provides a summary of recently published estimates of the rate of pediatric bipolar spectrum diagnoses in different clinical settings. The table also includes information about the demography of participants and the diagnostic methodology used. Most of the studies have limitations that make them suboptimal for application in at least some clinical settings, but they at least provide a starting point for establishing risk of PBD

TABLE 1
Strengths and Limitations of Using Local Estimates of Base Rates of Disorder

Strengths
Rates definitely will be pertinent to clinical setting of interest
Rates reflect potential moderators of base rate
Demographic characteristics
Referral practices
Pattern of recruitment and retention in treatment
Diagnostic rates will correspond highly with clinical diagnoses (as distinct from research diagnoses)
Limitations
Local rates may not be readily available (if no computer database of diagnoses)
Diagnostic practices may very widely across settings or even across clinicians within setting
Local clinical diagnoses have unknown interrater or retest reliability, although typically lower than research diagnoses based on structured or semistructured interviews
Local clinical diagnoses may underestimate rates of comorbidity and total number of diagnoses
Local diagnoses may be influenced by factors in addition to concern for accurate labeling
Concerns about reimbursement for services
Concerns about stigma associated with certain diagnoses
Concerns about labeling child with a chronic condition (label could persist in medical record after condition has remitted)

TABLE 2
Base Rates of Pediatric Bipolar Disorder in Different Clinical Settings

Setting	Base Rate	Demography	Diagnostic Method	Ref.
High school epidemiological	0.6%	Northwestern high school	K-SADS-PL	Lewinsohn et al., 2000 (Oregon Longitudinal Study of Depression)
Epidemiological	~0%	Southeastern	CAPA[p,y]	Costello et al., 1996 (Great Smoky Mountains Study of Youth, epidemiological)
Epidemiological	<2%	Midwest	Structured interview[p,y]	Kashani et al., 1987 (N = 150, ages 14–16 yr)
Community mental health center	5.9%	Midwestern urban, 48% nonwhite (42% black)	Clinical interview and treatment[p,y]	Youngstrom et al., in press b (N = 3,086, ages 4–18 yr)
General outpatient clinic	6% to 8%	Urban academic research centers	Washington University K-SADS[p,y]	Geller et al., 2002 (TEAM Study)
County wards (DCFS)	11%	Illinois	Clinical interview and treatment[y]	Naylor et al., 2002
Specialty ADHD outpatient service	15–17%	New England	K-SADS[p,y (only p young)]	Biederman et al., 1996
Incarcerated adolescents	2%	Midwestern urban (Chicago)	DISC[y]	Teplin et al., 2002
Incarcerated adolescents	22%	Texas	DISC[y]	Pliszka et al., 2000
Inpatient service	30% manic symptoms, <2% strict bipolar I	New York State	DICA, K-SADS[p,y]	Carlson and Youngstrom, 2003

Note: Articles similar to those comprising this table could be found by searching for "bipolar disorder" and "prevalence" or "epidemiology" in Medline, *PsycINFO*, or similar databases. K-SADS-PL = Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Life time; CAPA = Child and Adolescent Psychiatric Assessment; DCFS = Department of Child and Family Services; ADHD = Attention-Deficit/Hyperactivity Disorder; DISC = Diagnostic Interview Schedule for Children; DICA = Diagnostic Interview for Children and Adolescents.

[p] Parent interviewed as component of diagnostic assessment.

[y] Youth interviewed as part of assessment.

diagnoses. Given the available information, your colleague decides to begin with a base rate of 6%, roughly what has been reported in community mental health centers based on clinical diagnoses as well as the rate observed in outpatient academic medical centers using comprehensive semistructured interviews.

FAMILIAL RISK

The next step in the assessment process is to determine how much weight to assign to the family history of bipolar disorder. Although bipolar disorder is highly heritable, it is important to diagnose the child based on risk and symptoms versus "diagnosing the family." A *Medline* search combining the terms "bipolar disorder," "offspring," "adolescents OR children OR pediatric" and "children-of-impaired-parent" (an exact MeSH heading, offered by the *Medline* thesaurus) using the AND operation (narrowing down the search) revealed 38 hits. The most recently published meta-analysis also is the most exhaustive, so it seems a reasonable choice for the best current evidence (Hodgins et al., 2002). Its results indicate that children with a biological parent diagnosed with bipolar disorder average a fivefold increase in the likelihood of having a bipolar diagnosis themselves. Based on the genetics of family relationships, the fivefold increase in risk would be equally appropriate if bipolar disorder were identified in any first-degree relative (i.e., biological mother, biological father, full biological siblings). If bipolar illness is identified in the pedigree, but one step further removed (e.g., in a biological grandparent, aunt, uncle, half sibling), then the associated risk would be estimated as 2.5, half the risk associated with a first degree relative, because this set of relatives on average shares half as many genes as would a pair of first-degree relatives. In practical terms, changes in odds of 2.5 or less will not have much formal

impact on the final probability that a particular patient has PBD. Although family history is clinically valuable (particularly because it might help shed light on treatment history and potential responses to intervention), information about more distant relatives often will not be available and probably would not have a big impact on differential diagnosis even if gathered. Another recent meta-analysis indicates that no other risk factors besides family history have been sufficiently documented to justify their integration into clinical decision making (Tsuchiya et al., 2003).

COMBINING INFORMATION ABOUT FAMILY RISK AND BASE RATE

How much does the family history change the odds that this specific child has a PBD? The most accurate way of combining information about risk is through the use of Bayes' theorem. To do this involves algebra and multiplication, hampering the use of Bayes' theorem in practice. For this reason, many experts advocate the use of a "nomogram," which is a simple figure that allows the clinician to combine the starting probability of PBD (in this instance, the PBD base rate of 6%) with the change in risk (in this scenario, a likelihood ratio of 5, due to the biological father having bipolar disorder) (Guyatt and Rennie, 2002). After locating these two points on the nomogram, the clinician simply connects the dots and reads the new probability of PBD from the third column (Fig. 1). Using the nomogram with these values (6% base rate, 5.0 likelihood ratio) yields a posterior probability of 25%. This is quite close to the Bayesian exact estimate of 24.2% (see Guyatt and Rennie, 2002 for details).

Several aspects are worth noting about the nomogram and this approach to combining diagnostic information. One is that probability does not behave in a linear way. The scales of the nomogram are not marked off in a linear manner, nor is the revised/posterior probability linearly related to the starting/prior probability. Likelihood ratios change the *odds* of a diagnosis in a straightforward manner (multiplying prior odds by the likelihood ratio yields posterior odds), but most people are not familiar with odds and cannot interpret them as readily as probabilities. The nomogram allows the clinician to work directly with probabilities (accomplishing the transformation into odds and from odds back into probability) without requiring any mathematical computation.

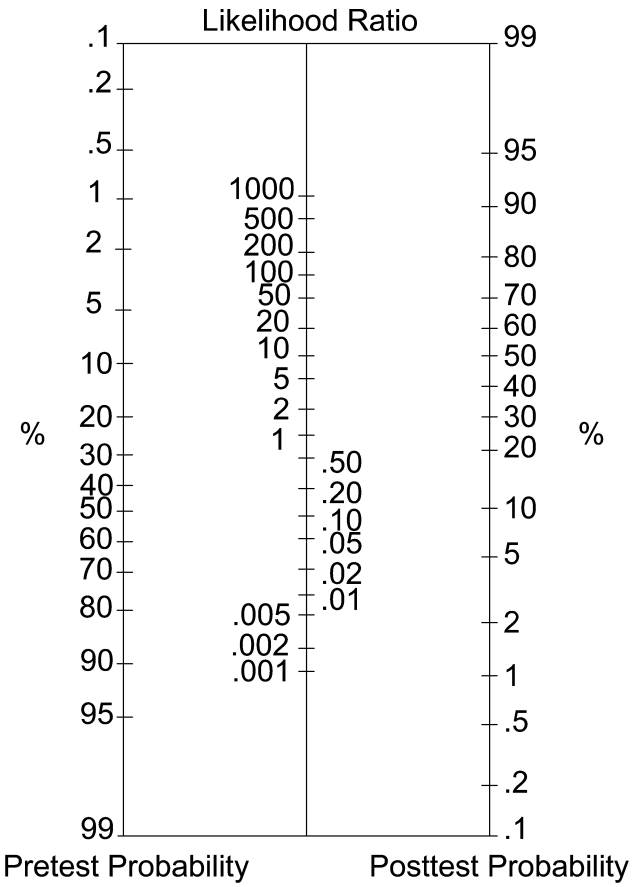

Fig. 1 Nomogram for combining probability and likelihood ratio.

A second observation is that the same piece of diagnostic information may produce substantially different posterior probabilities in different settings. This is driven by the differences in the base rate of the condition in different settings. For instance, assuming base rates of 1% or 30% (Table 2) would mean that the revised probability that a child has PBD after learning that he or she has a parent diagnosed with the condition would vary from 5% to 68%. Put another way, almost all the children ascertained in a nonclinical community setting (such as a public school) who have a parent with a bipolar disorder will not themselves currently have a PBD. At the other extreme, two of three youths on an inpatient unit where 30% of the patients have bipolar diagnoses are likely to also have a bipolar diagnosis if it has been diagnosed in a parent.

It may seem counterintuitive that the same piece of diagnostic information (familial history of bipolar disorder) could generate such different risk estimates. This has been well known in the diagnostic efficiency

literature, where it is often described in terms of the "positive and negative predictive values" of a test depending on the base rate of the target condition (Kraemer, 1992). In fact, the probabilities read from the third column of the nomogram are actually "positive predictive values," which can be interpreted as either the probability that a specific individual has the target diagnosis or alternately as the frequency with which people showing that test result at that particular clinical setting would have the condition (Kraemer, 1992). For the present example, the posterior probability of 24% could be interpreted as signifying either a 24% risk that this specific patient has PBD or an indication that 24% of all patients presenting to this clinic with a first-degree relative with a bipolar disorder will have PBD. Some concrete examples may make the concept more intuitive: When fishing for trout, a tug on the line is much more likely to indicate that a trout has been hooked when fishing in a well-stocked stream (a high base rate situation) versus in a depleted or polluted stream (lower base rate) or the ocean (where a trout would be a freak occurrence, but not statistically impossible).

The last observation is that the posterior probabilities will rarely be decisive. The 24% risk of PBD suggests that it is unlikely that this child has PBD, but many parents and practitioners would agree that this risk is too high to rule out the diagnosis. The risk exceeds the test/no-test threshold, indicating that more assessment is needed (Guyatt and Rennie, 2002). Practitioners might decide that additional assessment is warranted until the risk is below 10% (or 5%), given that there are relatively inexpensive assessment options and the costs of failing to diagnose PBD may be high. Conversely, the 24% probability of PBD is below the treatment threshold, indicating that treatments specific to bipolar disorder probably should not be started without further information gathering (Guyatt and Rennie, 2002). The treatment threshold is determined by the clinician in consultation with the family, and it should weigh information about the costs and benefits of treatment for both accurately diagnosed and misdiagnosed cases. Pragmatically, differences in cost and benefit do not have much of an impact on diagnostic decision thresholds unless they vary by a factor of 10 or more (Kraemer, 1992), suggesting that decision-making frameworks can tolerate a fair amount of difference in subjective opinions or actual costs. In the absence of overwhelming evidence of efficacy or harm, it becomes most practical to simply discuss the probability with the patient and guardian and decide on the assessment and treatment strategy taking personal preference into account.

Your colleague concludes that the 24% risk of PBD in her patient is above the test threshold, so she decides to search for tools that might help assess PBD. The next commentary details how to search for such instruments and how to incorporate them into an EBP framework.

Disclosure: Dr. Youngstrom is co-investigator on investigator-initiated research grants sponsored by Abbott and AstraZeneca Pharmaceuticals; he is the statistical expert for both protocols. He also consults with Otsuka Pharmaceuticals about assessment of pediatric bipolar disorder. Ms. Duax has no financial relationships to disclose.

REFERENCES

Biederman J, Faraone S, Mick E et al. (1996), Attention-deficit hyperactivity disorder and juvenile mania: an overlooked comorbidity? *J Am Acad Child Adolesc Psychiatry* 35:997–1008

Carlson G (2003), The bottom line. *J Child Adolesc Psychopharmacol* 13: 115–118

Carlson GA, Youngstrom EA (2003), Clinical implications of pervasive manic symptoms in children. *Biol Psychiatry* 53:1050–1058

Costello EJ, Angold A, Burns BJ et al. (1996), The Great Smoky Mountains Study of Youth: goals, design, methods, and the prevalence of DSM-III-R disorders. *Arch Gen Psychiatry* 53:1129–1136

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

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

Guyatt GH, Rennie D, eds. (2002), *Users' Guide to the Medical Literature.* Chicago: AMA Press

Hodgins S, Faucher B, Zarac A, Ellenbogen M (2002), Children of parents with bipolar disorder. A population at high risk for major affective disorders. *Child Adolesc Psychiatr Clin N Am* 11:533–553

Kashani JH, Carlson GA, Beck NC et al. (1987), Depression, depressive symptoms, and depressed mood among a community sample of adolescents. *Am J Psychiatry* 144:931–934

Kessler RC, Avenevoli S, Merikangas KR (2001), Mood disorders in children and adolescents: an epidemiologic perspective. *Biol Psychiatry* 49:1002–1014

Kessler RC, Rubinow DR, Holmes C, Abelson JM, Zhao S (1997), The epidemiology of DSM-III-R bipolar I disorder in a general population survey. *Psychol Med* 27:1079–1089

Kraemer HC (1992), *Evaluating Medical Tests: Objective and Quantitative Guidelines.* Newbury Park, CA: Sage

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

Lewinsohn PM, Klein DN, Seeley J (2000), Bipolar disorder during adolescence and young adulthood in a community sample. *Bipolar Disord* 2:281–293

Naylor MW, Anderson TR, Kruesi MJ, Stoewe M (2002), 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, San Francisco, October

Pliszka SR, Sherman JO, Barrow MV, Irick S (2000), Affective disorder in juvenile offenders: a preliminary study. *Am J Psychiatry* 157:130–132

Teplin LA, Abram KM, McClelland GM, Dulcan MK, Mericle AA (2002), Psychiatric disorders in youth in juvenile detention. *Arch Gen Psychiatry* 59:1133–1143

Tsuchiya KJ, Byrne M, Mortensen PB (2003), Risk factors in relation to an emergence of bipolar disorder: a systematic review. *Bipolar Disord* 5: 231–242

Weckerly J (2002), Pediatric bipolar mood disorder. *J Dev Behav Pediatr* 23:42–56

Weller EB, Danielyan AK, Weller RA (2004), Somatic treatment of bipolar disorder in children and adolescents. *Psychiatr Clin North Am* 27: 155–178

Youngstrom EA, Findling RL, Calabrese JR (2004), Effects of adolescent manic symptoms on agreement between youth, parent, and teacher ratings of behavior problems. *J Affect Disord* 82S:S5–S16

Youngstrom EA, Findling RL, Youngstrom JK, Calabrese JR (in press a), Towards an evidence-based assessment of pediatric bipolar disorder. *J Clin Child Adolesc Psychol*

Youngstrom EA, Kogos Youngstrom J, Starr M (in press b), Bipolar diagnoses in community mental health: Achenbach CBCL profiles and patterns of comorbidity. *Biol Psychiatry*