Notice of correction 4/12/2013:
The funding/support statement has been corrected to delete the reference to NIMH grant 1R01MH092450.

Irritability and Elation in a Large Bipolar Youth Sample: Relative Symptom Severity and Clinical Outcomes Over 4 Years

Jeffrey I. Hunt, MD; Brady G. Case, MD; Boris Birmaher, MD; Robert L. Stout, PhD; Daniel P. Dickstein, MD; Shirley Yen, PhD; Tina R. Goldstein, PhD; Benjamin I. Goldstein, MD, PhD; David A. Axelson, MD; Heather Hower, MSW; Michael Strober, PhD; Neal Ryan, MD; Lance Swenson, PhD; David R. Topor, PhD; Mary Kay Gill, MSN; Lauren M. Weinstock, PhD; and Martin B. Keller, MD

ABSTRACT

Objective: To assess whether relative severity of irritability symptoms versus elation symptoms in mania is stable and predicts subsequent illness course in youth with *DSM-IV* bipolar I or II disorder or operationally defined bipolar disorder not otherwise specified.

Method: Investigators used the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children to assess the most severe lifetime manic episode in bipolar youth aged 7–17 years who were recruited from 2000 to 2006 as part of the Course and Outcomes of Bipolar Youth prospective cohort study (N = 361), conducted at university-affiliated mental health clinics. Subjects with at least 4 years of follow-up (N = 309) were categorized as irritable-only (n = 30), elated-only (n = 42), or both irritable and elated (n = 237) at baseline. Stability of this categorization over follow-up was the primary outcome. The course of mood symptoms and episodes, risk of suicide attempt, and functioning over follow-up were also compared between baseline groups.

Results: Most subjects experienced both irritability and elation during follow-up, and agreement between baseline and follow-up group assignment did not exceed that expected by chance ($\kappa = 0.03$; 95% CI, −0.06 to 0.12). Elated-only subjects were most likely to report the absence of both irritability and elation symptoms at every follow-up assessment (35.7%, versus 26.7% of irritable-only subjects and 16.9% of those with both irritability and elation; $P = .01$). Baseline groups experienced mania or hypomania for a similar proportion of the follow-up period, but irritable-only subjects experienced depression for a greater proportion of the follow-up period than did subjects who were both irritable and elated (53.9% versus 39.7%, respectively; $P = .01$). The groups did not otherwise differ by course of mood episode duration, polarity, bipolar diagnostic type, suicide attempt risk, or functional impairment.

Conclusions: Most bipolar youth eventually experienced both irritability and elation irrespective of history. Irritable-only youth were at similar risk for mania but at greater risk for depression compared with elated-only youth and youth who had both irritability and elation symptoms.

J Clin Psychiatry 2013;74(1):e110–e117

Submitted: *May 1, 2012; accepted November 7, 2012 (doi:10.4088/JCP.12m07874).*
Corresponding author: *Jeffrey I. Hunt, MD, Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, 1011 Veterans Memorial Parkway, East Providence, RI 02895 (Jeffrey_Hunt@Brown.edu).*

Diagnosing pediatric bipolar disorder on the basis of a core mood symptom of irritability without elation is controversial, and it remains unclear how well the chief complaint of irritability serves as a marker for risk of developing bipolar disorder.[1–4] One study[5] of a large community sample found that irritability in adolescence predicted incidence of major depressive disorder, generalized anxiety disorder, and dysthymia, but not bipolar disorder, at adult follow-up. A clinical sample with severe mood dysregulation—chronic irritability without elation, pervasive negative mood, and attention-deficit/hyperactivity disorder–like symptoms of hyperactivity—demonstrated far lower risk of manic or mixed episodes than a sample of youth with bipolar disorder.[6] By contrast, other work has suggested an association between irritability and development of bipolar disorder, but only for those youth with episodic, rather than chronic, irritability. A study[7] of a community sample of approximately 700 children with irritability found that episodically irritable subjects were more likely than nonepisodically irritable subjects to experience a manic episode over a 3-year follow-up. Longitudinal studies[8–10] of youth with oppositional defiant disorder have similarly demonstrated a strong association between episodic irritability and subsequent diagnosis of mania.

Because most longitudinal assessments of a putative association between irritable mood and bipolar disorder in youth have not evaluated episodic irritability in stratified or moderator analyses, these data do not inform the question of whether episodic irritability without elation represents a manifestation of bipolar disorder. We previously examined whether subjects included for criterion A manic symptoms of episodic irritability without elation ("irritable-only," representing 10% of the sample) differed from others ("elated-only" and "both irritable and elated," representing 15% and 75%, respectively) in terms of baseline sociodemographic, phenomenological, and familial features.[2,11] No between-group differences in the bipolar disorder subtype, probability of psychiatric comorbidity, illness severity or duration, and family history of mania or other psychiatric disorders were found, with the exception of depression and alcohol abuse occurring more frequently in the irritable-only group.

We now examine the stability of irritability and elation symptoms and contrast the longitudinal course of irritable-only, elated-only, and both irritable and elated youth over 4 years using the previously established baseline *DSM-IV* criterion A grouping. Further, we assess whether irritable-only

Reprinted with correction to page e116.

Clinical Points

- Compared to bipolar youth with prominent elation symptoms, youth who are diagnosed with bipolar disorder on the basis of episodic irritability experience a similar clinical course but may be at greater risk for depression.
- Future research on irritability and elation in pediatric bipolar disorder would be aided by more precise assessment of covariation in irritability and elation during and between major mood episodes.
- Improved characterization of distinct illness courses may enhance the identification and utility of emerging genetic and neuroimaging markers and facilitate development of targeted treatments.

youth differ from the other groups in the polarity, severity, and duration of mood episodes, risk of suicide attempts, and global functional impairment. We hypothesized that the criterion A groups would not remain stable over time and that the course of these criterion A baseline groups would not differ and that these findings would provide support for the validity of episodic irritability without elation as a criterion A symptom of pediatric bipolar disorder.

METHOD

Subjects

Subjects were drawn from the Course and Outcomes of Bipolar Youth (COBY) prospective cohort study, conducted at university-affiliated mental health clinics; the study is described elsewhere.[2,11,12] At study entry, the subjects (recruited from 2000 to 2006) were aged 7 years and 0 months to 17 years and 11 months and met either *DSM-IV* criteria for bipolar I or II disorder or COBY criteria for bipolar disorder not otherwise specified (NOS). Bipolar disorder NOS was defined as the presence of elated mood plus 2 associated manic symptoms, or irritable mood plus 3 *DSM-IV*–associated manic symptoms, along with a change in the level of functioning, with a duration of a minimum of 4 hours within a 24-hour period and at least 4 cumulative lifetime days meeting the criteria.[12] The current analyses are limited to subjects administered baseline assessment of the most serious lifetime manic episode and at least 4 years of follow-up assessment after study entry (N = 309).

Informed Consent

Institutional review boards at each site approved the study protocol before subject enrollment, and procedures were the same at each site. Informed consent was obtained from each subject's parent/guardian and from subjects aged 14 years or older. For younger subjects, study procedures were explained in age-appropriate language, and verbal assent was obtained.

Interview Methods

A trained research clinician assessed each youth subject and a parent or primary caregiver using a semistructured interview. Principal investigators reviewed all available information before reaching a consensus diagnosis on intake and follow-up ratings. Interviewers were not blind to subjects' prior diagnoses. Interrater reliability was evaluated using audiotapes of randomly selected interviews and is reported below by assessment.

Baseline Diagnostic Assessment

Mood symptoms, episodes, and disorder. Mood symptoms were assessed using the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present Episode (K-SADS-P)[13] and additional items from the K-SADS Mania Rating Scale (K-SADS-MRS).[14] The reliability of intake bipolar diagnosis presence and type was $\kappa = 0.74$.[12]

Mood symptom severity was determined for the most severe week in the month prior to intake assessment using the 12-item K-SADS-P depression rating score and the K-SADS-MRS. Symptom severity for the most severe week in the subject's lifetime was rated using the K-SADS Present and Lifetime Version (K-SADS-PL)[15] with the first 84 enrolled subjects and the K-SADS-MRS and K-SADS-PL mood disorder sections for subsequent subjects. Symptoms were rated from 1 (normal or not present) to 6 (extreme symptoms).

Age at onset was defined as the initiation of clinically significant mood symptoms that affected the child's functioning, and the first and the most recent episode of each type of major mood episode was recorded.

DSM-IV *criterion A symptom grouping at baseline.* Subjects (N = 361) were classified into 1 of 3 criterion A mood symptom groups using the most serious lifetime manic episode scores on the K-SADS-MRS: irritable-only (Irritable, n = 36), elated-only (Elated, n = 54), or both elated and irritable (Both, n = 271). Mild or greater severity established the presence of elation or irritability. All subjects met threshold criteria for elation and/or irritability at intake or at the most serious past episode, defined as mild or greater severity (score ≥ 3). The K-SADS-MRS is a 15-item inventory with excellent interrater reliability (intraclass correlation coefficient = 0.97) and convergent validity with the Clinical Global Impressions-Severity of Illness scale ($r_s = 0.91$), and it differentiates bipolar patients from those without significant manic symptoms.[14] Items on the K-SADS-MRS are derived from the K-SADS-P 1986 version and include a mood lability item. Symptoms are rated from 1 (normal or not present) to 6 (extreme symptoms). Mild elation (rated as 3 on this scale) reflects the presence of a "definitely elevated mood and optimistic outlook that is somewhat out of proportion to his/her circumstances." Mild irritability (rated as 3 on this scale) reflects that the patient "often (at least 3 separate times for at least 3 hours each week) feels definitely more angry and irritable than called for by the situation, relatively frequent but never very intense. Also is often argumentative, quick to express annoyance."

Other baseline assessments. Psychiatric disorders other than mood disorders were assessed using the K-SADS-PL.

The Children's Global Assessment Scale (CGAS)[16] was used to establish global level of functioning. Demographic and clinical characteristics of baseline criterion A groups were published previously[11] (see Table 1).

Longitudinal Follow-Up

Timing and frequency of follow-up assessments. Follow-up data on the first 4 years (208 weeks) after enrollment was included in the current analyses. These data were drawn from subjects with a mean follow-up period of 259 weeks (standard deviation = 100 weeks; median = 267 weeks; range, 26–434 weeks), with a mean of 7.4 assessments (range, 1–15 assessments) and a mean interval between assessments of 8.2 months.

Mood symptoms. The K-SADS-MRS and K-SADS-P depression scales were used to assess manic symptoms, including irritability and elation, for the most symptomatic week in the month preceding follow-up assessment for all subjects. Subjects were not necessarily within a manic or depressed episode during scheduled follow-up assessments.

Mood episodes and disorder. Weekly change in mood symptomatology was assessed over time using the Psychiatric Status Rating (PSR) scales from the semistructured adolescent version of the Longitudinal Interval Follow-up Evaluation (A-LIFE).[17,18] The A-LIFE PSR evaluates symptom course by identifying change points, frequently anchored by memorable dates for the subject (eg, holidays, beginning of school). Subjects are queried regarding their mood symptoms since the last interview, and then this information is translated into ratings for each week of the interim period. Rating values are operationally linked to the *DSM-IV* criteria and indicate symptom severity and impairment. The A-LIFE PSR mood scores range from 1 for no symptoms, 2–4 for increasing symptom severity and impairment that does not meet *DSM-IV* episode criteria ("subthreshold"), and 5–6 for increasing severity or impairment meeting *DSM-IV* criteria. Consensus scores obtained after interviewing parents and their children were used for the analyses. The week of mood symptoms onset and offset were recorded, providing information on episode duration and the percentage of time syndromal and subsyndromal for specific mood states. Interrater reliability for rating manic, mixed, or hypomanic episodes (using the A-LIFE PSR) was $\kappa = 0.62$ and, for follow-up major depressive episodes, was also $\kappa = 0.62$. The Kendall *W* statistic for percentage of follow-up weeks in euthymic, full syndromal, and subsyndromal mood states was ≥ 0.75.[19]

DSM-IV *criterion A symptom grouping at follow-up.* Group assignment during follow-up was determined in a similar fashion to baseline assignment, except that maximum (ie, most severe over the follow-up period) K-SADS-MRS scores were drawn from the 4-year follow-up period. Subjects were classified as being neither irritable nor elated at follow-up if all follow-up K-SADS-MRS irritability and elation scores were less than 3 (indicating "none or minimal" severity).

Suicide attempt. Suicide attempts were identified using the subject or parent report on the K-SADS depression module and Summary Lifetime Diagnostic Checklist suicide attempt item. A past or present suicide attempt was defined as any self-injurious act that reached or exceeded an operationalized threshold of lethal intent and/or medical lethality.

Study Outcomes

We examined the stability of criterion A group assignments and of irritability and elation symptom severity at baseline versus follow-up. We also tested for criterion A group differences in the follow-up period using the following measures: (1) severity of criterion A and B manic symptoms, (2) proportion of time syndromal for mania (A-LIFE PSR ≥ 5) and syndromal or subsyndromal (A-LIFE PSR ≥ 3) for hypomania and for depression, (3) time to relapse and recovery from manic and depressive episodes, (4) bipolar subtype at the conclusion of follow-up, (5) probability of 1 or more suicide attempts, and (6) change in CGAS.

Statistical Analyses

All *P* values are from 2-tailed tests, with $\alpha = .05$. Analyses were conducted with SAS software, version 9.2 (SAS Institute; Cary, North Carolina).

Differences between criterion A groups at baseline. Group differences in categorical characteristics were tested with a χ^2 statistic, while differences in continuous characteristics were tested with an *F* statistic for normally distributed values or the Kruskal-Wallis test for non-normally distributed values.[11]

Criterion A symptom stability. Stability of criterion A irritability and elation symptoms was assessed using 3 approaches. First, to assess stability of criterion A group assignment, we calculated κ, a measure of concordance beyond that expected by chance, for agreement between the criterion A group at baseline and at follow-up. Subjects who were neither irritable nor elated at follow-up were excluded. Second, to assess the stability of continuous measures of irritability and elation symptoms, we calculated Pearson correlations of maximum irritability and elation according to K-SADS-MRS scores at baseline versus follow-up. Finally, to assess whether baseline category predicted continuous measures of criterion A symptom severity during follow-up, we conducted separate repeated-measures hierarchical linear model (HLM) analyses,[20] with severity of irritability and elation as the outcomes. Each model was fit using a linear time effect, baseline group, and a time-by-baseline group interaction term. Groups were considered to differ if the time-by-baseline group interaction term was significant (indicating a differential rate of change in severity over follow-up) or, in the absence of a significant interaction, if the baseline group predicted severity (indicating a differential level of severity across follow-up).

Other between-group contrasts of clinical and functional course. Between-group differences in follow-up clinical and functional scores, percentage of time ill, and probability of suicide attempt were tested using HLM

Reprinted with correction to page e116.

Table 1. Sociodemographic and Clinical Characteristics of Bipolar Youth at Baseline, by Baseline Criterion A Group Assignment (N = 361)

Characteristic	Irritable Only (n = 36)	Elated Only (n = 54)	Both (n = 271)	Group Contrast *P* Value[a]
Sociodemographic characteristics				
Age, mean ± SD, y	10.5 ± 2.8[b]	12.7 ± 3.4[c]	12.7 ± 3.3[c]	**.01**[d]
Gender, male, %	44.4	42.6	53.1	.3[e]
Race, white, %	88.9	72.2	79.7	.2[e]
Ethnicity, Hispanic, %	2.8	7.4	7.0	.8[e]
Socioeconomic status, mean ± SD[f]	3.1 ± 1.2	3.6 ± 1.2	3.3 ± 1.2	.2[d]
Living with both natural parents, %	41.7	37.0	41.3	.8[e]
Clinical characteristics				
Pubertal status category, %				**.03**[e]
I	37.9	19.0	28.5	
II–III	41.4	23.8	24.0	
IV–V	20.7	57.1[c]	47.5[c]	
Bipolar type, %				.1[e]
Bipolar I	44.4	57.4	66.1	
Bipolar II	8.3	5.6	5.2	
Bipolar not otherwise specified	47.2	37.0	28.8	
Age at onset of mood symptoms, mean ± SD, y	6.8 ± 3.4	8.3 ± 4.4	8.4 ± 4.2	.2[g]
Age at onset of bipolar disorder, mean ± SD, y	8.2 ± 3.3	9.6 ± 4.3	9.2 ± 4.0	.3[g]
Duration of bipolar disorder, mean ± SD, y	2.4 ± 1.7	3.1 ± 2.6	3.4 ± 2.7	.05[d]

[a]Significant *P* values (< .05) are shown in bold type.
[b]Group differs significantly from elated-only and both irritable and elated groups in pairwise contrasts.
[c]Group differs significantly from irritable-only group in pairwise contrasts.
[d]*P* value obtained from *F* test.
[e]*P* value obtained from χ^2 test.
[f]See Hunt et al[11] on measurement of socioeconomic status.
[g]*P* value obtained from Kruskal-Wallis test.

Table 2. Distribution of Baseline Versus Follow-Up Criterion A Group Assignment of Bipolar Youth Subjects (N = 246)[a]

Baseline Group	Irritable Only (n = 55), n (% of baseline group)	Elated Only (n = 29), n (% of baseline group)	Both (n = 162), n (% of baseline group)
Irritable only (n = 22)	10 (45)	0 (0)	12 (55)
Elated only (n = 27)	6 (22)	1 (4)	20 (74)
Both (n = 197)	39 (20)	28 (14)	130 (66)

[a]Distributions exclude subjects with neither irritability nor elation during follow-up (n = 63).

analyses as described above. Differences in time to relapse or recovery were tested in Cox proportional hazards regression models accommodating multiple events per participant using the marginal method.[21] Survival probabilities were described using Kaplan-Meier survival curves.

Testing for confounding of between-group contrasts by age and sex. To evaluate whether significant between-group clinical differences could be explained by group differences in age or sex, we assessed whether the effect of baseline group remained significant after adding age and sex as main effects to the HLM and Cox analyses described above. We also stratified between-group contrasts of criterion A group assignment at follow-up by a dichotomized age variable (≤ 11 vs ≥ 12 years) and by sex and assessed whether patterns observed in the total sample persisted in each stratum.

RESULTS

Baseline Criterion A Group Differences

Groups differed only by age distribution and pubertal status at baseline; the Irritable group was significantly younger and less physically developed than the other groups (Table 1).

Stability of Criterion A Groups and Symptoms

Agreement of group assignment at baseline versus follow-up. Among those subjects who experienced clinically relevant criterion A symptoms, the κ coefficient of agreement between baseline group and follow-up group was κ = 0.03 (95% CI, −0.06 to 0.12), indicating poor agreement that did not differ statistically from that expected by chance. Among these subjects, the rate of agreement for baseline versus follow-up was 33% for Irritable, 2% for Elated, and 55% for Both. Table 2 shows the proportion of the sample with specific patterns of agreement and disagreement between baseline and follow-up group assignment. Most subjects in the Irritable and Elated baseline groups were in the Both group at follow-up.

Correlation of baseline versus follow-up symptom severity. The maximum irritability score during follow-up was modestly correlated with the maximum baseline score ($r = 0.15$, $P = .008$), while the maximum elation score during follow-up was not significantly correlated with the corresponding baseline score ($r = 0.06$, $P = .32$).

Between-Group Contrasts During Follow-Up

All between-group contrasts are shown in Table 3 except for contrasts of time to mood episode relapse and

Reprinted with correction to page e116.

Table 3. Clinical and Functional Outcomes at Follow-Up, by Baseline Category A Group Assignment (N = 309)

Follow-Up Outcome Measure	Irritable Only (n = 30)	Elated Only (n = 42)	Both (n = 237)	Group Contrast *P* Value[a]
Category A symptom severity score, mean[b]				
Irritability	2.4	2.1	2.3	.16
Elation	1.7	2.0	2.2	.06
Subjects with neither irritability nor elation, %[c]	26.7	35.7	16.9	**.01**
Category B symptom severity score, mean[b]				
Grandiosity	1.7	1.4	1.8	**.006**
Decreased need for sleep	1.3	1.6	1.8	**.002**
Accelerated speech	2.2	1.9	2.3	.10
Racing thoughts	1.7	1.8	1.9	.92
Flight of ideas	1.8	1.8	2.1	.13
Distractibility	2.0	1.7	2.0	**.04**[d]
Motor hyperactivity	2.3	2.1	2.3	.44
Poor judgment	1.9	1.5	1.8	**.02**[d]
Unusual energy	2.0	2.0	2.2	.76
Increased goal-directed activity	1.6	1.6	1.7	.59
Uninhibited people seeking	1.4	1.3	1.4	.16
Increased productivity	1.2	1.4	1.4	.27
Increased creativity	1.5	1.4	1.5	.25
Hypersexuality	1.3	1.3	1.4	.68
Inappropriate laughing	1.6	1.6	1.7	.19
Percent time syndromal or subsyndromal, mean[b]				
Mania	4.3	3.3	4.5	.82
Hypomania	36.9	29.3	38.8	.10
Depression	53.9	43.0	39.7	**.01**
Bipolar diagnostic status at 48 months, %				.63
Bipolar I	61.5	62.9	72.6	
Bipolar II	11.5	17.1	9.5	
Bipolar not otherwise specified	26.9	20.0	17.9	
Subjects with suicide attempt, %	7.4	19.1	14.0	.40
CGAS score, mean[b]	61.1	61.8	58.9	.07

[a]Significant *P* values (< .05) are shown in bold type.
[b]Means shown are least-squares means obtained from generalized estimating equation models; *P* values are from model term for group main effect.
[c]*P* value is from χ^2 contrast between group probability distributions.
[d]Group contrast no longer significant at *P* < .05 after inclusion of age and sex as main effects in the model. See Method section for modeling procedures.
Abbreviation: CGAS = Children's Global Assessment Scale.

remission, which are reported in the text and illustrated in Figure 1.

Criterion A symptom severity. There were no significant between-group differences in criterion A symptom severity during follow-up (see Table 3).

All groups showed a significant decrease in severity of irritability and elation scores on the K-SADS-MRS over time (time effect, *P* < .001). Among all subjects, the irritability score fell from a mean of 2.8 at baseline to 1.9 at year 4, while the elation score fell from a mean of 2.4 to 1.7 over the same time period. A trend was observed for the mean maximum elation score to be higher in the Both group (2.2) than in the Irritable (1.7) and Elated (2.0) groups (*P* = .06).

Probability of no significant irritability or elation. Baseline groups differed in their probability of reporting neither clinically significant irritability nor clinically significant elation in the month preceding any follow-up assessment: 26.7% of the Irritable group, 35.7% of the Elated group, and 16.9% of the Both group reported neither symptom over the period of follow-up (*P* = .01). This pattern was observed in stratified analyses for female and school-aged subjects but not for male or adolescent subjects.

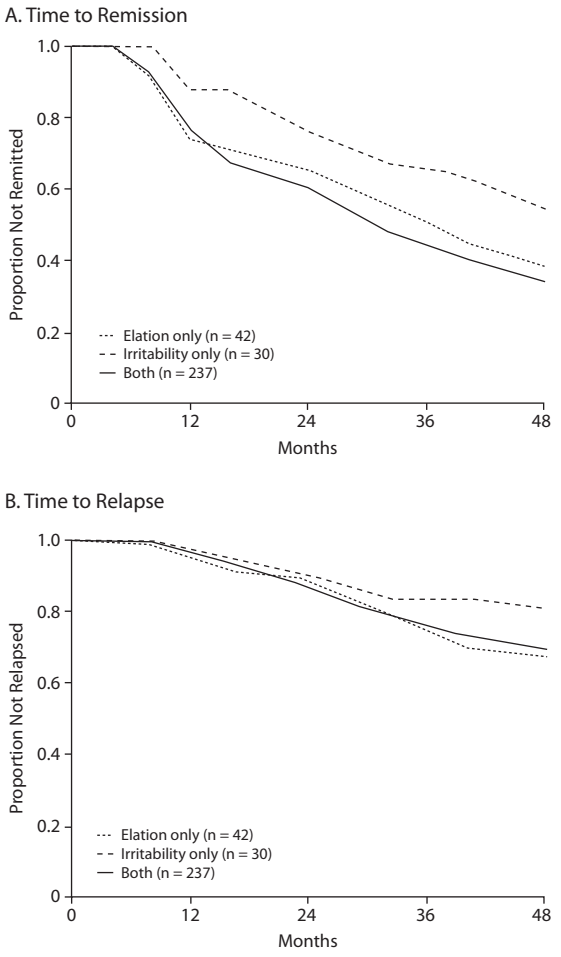
Criterion B symptom severity. Significant group differences were found in the mean severity score of 4 of the 16 K-SADS-MRS criterion B symptoms: decreased need for sleep (Both [1.8] greater than Elated [1.6] and Irritable [1.3]; *P* = .002), grandiosity (Elated [1.4] less than Both [1.8] and Irritable [1.7]; *P* = .006), poor judgment (Elated [1.5] less than Both [1.8] and Irritable [1.9]; *P* = .02), and distractibility (Both [2.0] and Irritable [2.0] greater than Elated [1.7]; *P* = .04). Observed group symptom differences were not meaningfully altered after adjustment for differences in age or sex except for symptoms of distractibility and poor judgment, which became nonsignificant.

Proportion of time syndromal or subsyndromal. Groups significantly differed by yearly percentage of time syndromal or subsyndromal for depression (*P* < .01). In pairwise contrasts, the Irritable group was more depressed more often than the Both group (least-squares mean probability of being syndromal or subsyndromal, 53.9% vs 39.7%, respectively; *P* = .01) but not the Elated group (53.9% vs 43.0%, respectively; *P* = .11). There were no between-group differences in yearly percentage of time syndromal or subsyndromal for mania or hypomania. Observed differences were not meaningfully altered after adjustment for group differences in age or sex.

Time to mood episode remission and relapse. There were no between-group differences in time to remission from a major depressive episode or a manic episode. There were also

Reprinted with correction to page e116.

Figure 1. Descriptive Kaplan-Meier Survival Curves of (A) Time to Remission From and (B) Time to Relapse to Major Depressive Episode, by Baseline Category A Group[a]



[a]No between-group differences were observed in Cox proportional hazards regressions. See Method section for modeling procedures.

no between-group differences in time to relapse for manic episode or major depressive episode. No group differences emerged after adjustment for age and sex. Descriptive Kaplan-Meier probability distributions for time to remission from and relapse to major depression are presented in Figure 1.

Change in bipolar diagnostic type. There were no between-group differences in the proportion of subjects who changed from bipolar disorder NOS to bipolar I or II disorder or from bipolar II disorder to bipolar I disorder ($P = .63$).

Suicide attempt. The proportion of subjects with 1 or more suicide attempts over the follow-up period did not significantly differ between groups, but the magnitudes or estimated group differences were large. There were no completed suicides in COBY subjects over the period of follow-up.

Global functioning. There were no between-group differences in CGAS scores. The mean score over the follow-up period was 60.7.

DISCUSSION

The main goals of this longitudinal follow-up study were to determine whether the relative severity of the *DSM-IV* criterion A manic symptoms of irritability versus elation in bipolar disorder was stable and distinguished illness course over 4 years. Our results were contradictory. We found that the relative severity of irritability versus elation was unstable. Maximum criterion A symptom severity at baseline and follow-up was modestly correlated for irritability and uncorrelated for elation, baseline group assignment failed to predict the course of irritability or elation severity, and baseline and follow-up assignments did not agree more than would be expected by chance. Together these findings suggest that the relative severity of irritability and elation are unstable and should not differentiate underlying clinical differences in bipolar disorder or meaningfully predict illness course.

Nonetheless, baseline group assignment based on relative symptom severity during the most serious lifetime manic episode did predict some differences in symptom severity and risk of mood episodes. Subjects categorized as irritable-only at baseline were syndromal or subsyndromal for depression for a greater proportion of the follow-up period than subjects categorized as both irritable and elated. While all groups experienced mania or hypomania for a similar proportion of the follow-up period, elated-only youth at baseline were most likely to report the absence of irritability and elation at every follow-up assessment. The scores on the K-SADS-MRS reflect that the subjects were not necessarily within a manic episode at the scheduled follow-up assessments.

We found statistically significant group differences in the mean severity of some criterion B symptoms, but the magnitude of these differences was too small to be of clear clinical significance.

No comparable studies of bipolar youth have reported on longitudinal course of criterion A symptoms, but data from adult samples have suggested that irritability confers distinctive risks. Overt irritability and psychomotor agitation appear to be markers for specific negative outcomes in adults in a subsyndromal bipolar mixed state,[22] such as greater risk for suicide attempts[23–25] and longer duration of episodes at intake.[23] We found no increase in rate of suicide attempts and no difference between the criterion A groups with regard to time to remission in the irritable-only group over the 4 years.

Irrespective of baseline group assignment, subjects in our sample were more commonly syndromal or subsyndromal for depression than for mania, a finding concordant with adult studies of bipolar disorder.[22,26] A 12-year follow-up of bipolar I disorder found that threshold and subthreshold depressive symptoms were over 3 times more frequent than threshold manic symptoms.[22] The same group found that, in bipolar II disorder, the symptomatic course is dominated by depressive rather than hypomanic or cycling/mixed episodes.[26]

An important study limitation is that the primary instrument used to longitudinally track weekly mania severity, the

Reprinted with correction to page e116.

A-LIFE PSR, does not assess elation and irritability separately. The K-SADS-MRS, which we did use to specifically assess irritability and elation, was used only for the most severe week in the month prior to each 6-month follow-up evaluation, which may or may not have been during a mood episode. By contrast, the baseline criterion A group assignments were derived using the most severe lifetime manic episode. Therefore, our baseline measures of criterion A symptom severity may be more clinically relevant than our follow-up measures. This difference may have contributed to the apparent longitudinal instability in relative criterion A symptom severity that we observed and also may explain why the baseline measures might predict depression course more powerfully than subsequent criterion A symptoms. Our baseline assessments were subject to recall bias, as were our follow-up assessments, although the interval between follow-up assessments was 8 months on average. We did not have data on interrater and test-retest reliability of our criterion A group assignments, and the reliability assessment procedure for bipolar diagnosis did not measure information variance (ie, independent interviewers may have elicited different information from the same subject), which could be an unknown source of diagnostic unreliability. The reliability of our weekly mood assessment, the A-LIFE PSR, is "good"[27] or "substantial"[28] by conventional interpretation, but measurement error of the A-LIFE PSR could have contributed to underestimation of group differences in illness course. Because the study sample was recruited by methods including referral from clinical programs and because the majority of subjects in COBY had already received treatment, our findings may not generalize well to never-treated youth.

The main clinical implication of this study is that, compared to bipolar youth with prominent elation symptoms, youth carefully diagnosed with bipolar disorder on the basis of episodic irritability experience a similar clinical course but may be at greater risk of depression. Future research on irritability and elation in pediatric bipolar disorder would be aided by more precise assessment of covariation in irritability and elation during and between major mood episodes. Improved characterization of distinct illness courses may enhance the identification and utility of emerging genetic and neuroimaging markers and facilitate development of targeted treatments.

Author affiliations: Department of Psychiatry and Human Behavior, Emma Pendleton Bradley Hospital (Drs Hunt, Case, and Dickstein) and Department of Psychiatry and Human Behavior, Butler Hospital (Drs Stout, Yen, Weinstock, and Keller and Ms Hower), Alpert Medical School, and Department of Health Services, Policy, and Practice (Dr Case), Brown University, Providence, Rhode Island; Western Psychiatric Institute and Clinic, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania (Drs Birmaher, T. R. Goldstein, B. I. Goldstein, Axelson, and Ryan and Ms Gill); Department of Psychiatry, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada (Dr B. I. Goldstein); Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles (Dr Strober); Department of Psychology, Suffolk University, Boston, Massachusetts (Dr Swenson); and Veterans Affairs Boston Healthcare System and Department of Psychiatry, Harvard Medical School, Boston, Massachusetts (Dr Topor).

Potential conflicts of interest: **Dr Hunt** has received an honorarium from Wiley Publishers. **Dr Case** has received research grant support from the American Academy of Child and Adolescent Psychiatry (AACAP) Eli Lilly Pilot Research Award and the American Psychiatric Association AstraZeneca Young Minds in Psychiatry Award and has been a consultant for Optum Behavioral Health and Blue Cross Blue Shield of Rhode Island. **Dr Birmaher** has received research grant support from the National Institute of Mental Health (NIMH) and has received publication royalties from Random House and Lippincott Williams & Wilkins. **Dr T. R. Goldstein** has received research grant support from NIMH, the National Institute on Drug Abuse (NIDA), and The Pittsburgh Foundation and has received publication royalties from Guilford Press. **Dr B. I. Goldstein** has received research grant support from Pfizer, has received honoraria from Purdue Pharma, and has been a consultant for Bristol-Myers Squibb. **Dr Keller** has received research grant support from Pfizer; has been a consultant for CeNeRx, Medtronic, and Sierra Neuropharmaceuticals; and has served on the advisory board for CeNeRx. **Drs Stout**, **Dickstein**, **Yen**, **Axelson**, **Strober**, **Ryan**, **Swenson**, **Topor**, and **Weinstock** and **Mss Hower** and **Gill** have no potential conflicts of interest to report.

Funding/support: Supported by NIMH grants MH59691 (Drs Hunt, Yen, Ryan, and Keller and Ms Hower), MH59929 (Drs Birmaher, Dickstein, T. R. Goldstein, B. I. Goldstein, and Axelson and Ms Gill), and MH59977 (Dr Strober). Dr Case was supported in part by AACAP-NIDA grant K12 DA000357-11. Dr Dickstein also received support from NIMH grants 5R01MH087513, 1R21MH096850, and 5R01MH087510.

Disclaimer: The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of NIMH or the National Institutes of Health.

Acknowledgments: The authors thank past Course and Outcomes of Bipolar Youth faculty, raters, and data personnel and those involved with the current study: ▪ Alpert Medical School of Brown University, Providence, Rhode Island: Christianne Esposito-Smythers, PhD (faculty); Christie Rizzo, PhD (faculty); Kerry Gagnon, BA (rater); Matthew Killam, BA (rater); and Marguerite Shashoua, BS (rater) ▪ University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania: M. Rasim Diler, MD (faculty); Soledad Romero, MD (faculty); Regina Sala, MD (faculty); Jessica Bell Wrona, BA (rater); Katie Aronson, BA (statistical support); Wonho Ha, MS (statistical support); Satish Iyengar, PhD (statistical support); Fangzi Liao, MS (statistical support); Vicky Tzanakos, MS (statistical support); Heather Kumar, BA (rater); and Sharon Nau, BA (rater) ▪ University of California Los Angeles: Norman Kim, PhD (faculty and rater), and Eunice Kim, PhD (faculty and rater). These acknowledged individuals have no potential conflicts of interest to report.

REFERENCES

1. Ghaemi SN, Bauer M, Cassidy F, et al; ISBD Diagnostic Guidelines Task Force. Diagnostic guidelines for bipolar disorder: a summary of the International Society for Bipolar Disorders Diagnostic Guidelines Task Force Report. *Bipolar Disord*. 2008;10(1, pt 2):117–128.
2. Birmaher B, Axelson D. Course and outcome of bipolar spectrum disorder in children and adolescents: a review of the existing literature. *Dev Psychopathol*. 2006;18(4):1023–1035.
3. Biederman J, Klein RG, Pine DS, et al. Resolved: mania is mistaken for ADHD in prepubertal children. *J Am Acad Child Adolesc Psychiatry*. 1998;37(10):1091–1096.
4. Carlson GA. Early onset bipolar disorder: clinical and research considerations. *J Clin Child Adolesc Psychol*. 2005;34(2):333–343.
5. Stringaris A, Cohen P, Pine DS, et al. Adult outcomes of youth irritability: a 20-year prospective community-based study. *Am J Psychiatry*. 2009;166(9):1048–1054.
6. Stringaris A, Baroni A, Haimm C, et al. Pediatric bipolar disorder versus severe mood dysregulation: risk for manic episodes on follow-up. *J Am Acad Child Adolesc Psychiatry*. 2010;49(4):397–405.
7. Leibenluft E, Cohen P, Gorrindo T, et al. Chronic versus episodic irritability in youth: a community-based, longitudinal study of clinical and diagnostic associations. *J Child Adolesc Psychopharmacol*. 2006;16(4):456–466.
8. Nock MK, Kazdin AE, Hiripi E, et al. Lifetime prevalence, correlates, and persistence of oppositional defiant disorder: results from the National Comorbidity Survey Replication. *J Child Psychol Psychiatry*. 2007;48(7):703–713.
9. Biederman J, Petty CR, Dolan C, et al. The long-term longitudinal course of oppositional defiant disorder and conduct disorder in ADHD boys: findings from a controlled 10-year prospective longitudinal follow-up study. *Psychol Med*. 2008;38(7):1027–1036.
10. Kim-Cohen J, Caspi A, Moffitt TE, et al. Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Arch Gen Psychiatry*. 2003;60(7):709–717.
11. Hunt J, Birmaher B, Leonard H, et al. Irritability without elation in a large

Reprinted with correction to page e116.

bipolar youth sample: frequency and clinical description. *J Am Acad Child Adolesc Psychiatry*. 2009;48(7):730–739.
12. Axelson D, Birmaher B, Strober M, et al. Phenomenology of children and adolescents with bipolar spectrum disorders. *Arch Gen Psychiatry*. 2006;63(10):1139–1148.
13. Chambers WJ, Puig-Antich J, Hirsch M, et al. The assessment of affective disorders in children and adolescents by semistructured interview: test-retest reliability of the Schedule for Affective Disorders and Schizophrenia for School-Age Children, Present Episode Version. *Arch Gen Psychiatry*. 1985;42(7):696–702.
14. Axelson D, Birmaher BJ, Brent D, 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(4):463–470.
15. Kaufman J, Birmaher B, Brent D, 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(7):980–988.
16. Shaffer D, Gould MS, Brasic J, et al. A Children's Global Assessment Scale (CGAS). *Arch Gen Psychiatry*. 1983;40(11):1228–1231.
17. Keller MB, Lavori PW, Friedman B, et al. The Longitudinal Interval Follow-up Evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry*. 1987;44(6):540–548.
18. Warshaw MG, Dyck I, Allsworth J, et al. Maintaining reliability in a long-term psychiatric study: an ongoing inter-rater reliability monitoring program using the Longitudinal Interval Follow-up Evaluation. *J Psychiatr Res*. 2001;35(5):297–305.
19. Axelson DA, Birmaher B, Strober MA, et al. Course of subthreshold bipolar disorder in youth: diagnostic progression from bipolar disorder not otherwise specified. *J Am Acad Child Adolesc Psychiatry*. 2011;50(10):1001–1016.
20. Raudenbush SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Thousand Oaks, CA: Sage Publications; 2002.
21. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distribution. *J Am Stat Assoc*. 1989;84(408):1065–1073.
22. Judd LL, Akiskal HS, Schettler PJ, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch Gen Psychiatry*. 2002;59(6):530–537.
23. Maj M, Pirozzi R, Magliano L, et al. Agitated depression in bipolar I disorder: prevalence, phenomenology, and outcome. *Am J Psychiatry*. 2003;160(12):2134–2140.
24. Balázs J, Benazzi F, Rihmer Z, et al. The close link between suicide attempts and mixed (bipolar) depression: implications for suicide prevention. *J Affect Disord*. 2006;91(2–3):133–138.
25. Swann AC, Moeller FG, Steinberg JL, et al. Manic symptoms and impulsivity during bipolar depressive episodes. *Bipolar Disord*. 2007;9(3):206–212.
26. Judd LL, Akiskal HS. Depressive episodes and symptoms dominate the longitudinal course of bipolar disorder. *Curr Psychiatry Rep*. 2003;5(6):417–418.
27. Fleiss JL. *Statistical Methods For Rates And Proportions*. 2nd ed. New York, NY: John Wiley; 1981.
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.

Editor's Note: We encourage authors to submit papers for consideration as a part of our Focus on Childhood and Adolescent Mental Health section. Please contact Karen D. Wagner, MD, PhD, at kwagner@psychiatrist.com.

Reprinted with correction to page e116.
J Clin Psychiatry 74:1, January 2013