Supplemental Methods

Scales Administered at Baseline (in alphabetical order):

Child Affective Lability Scale (CALS)[1,2]: Developed to assess mood lability in youth. We used subscales identified by a previous factor analysis in this population: Irritability, Mania, and Anxiety/Depression. Parent- and child-report; no time period is specified.

Child Behavioral Checklist (CBCL)[3]: A standardized instrument that assesses a wide range of parent-reported psychopathology in youth aged 4-18 years. Includes the following subscales: Anxious/Depressed, Withdrawn, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior. Symptoms assessed over the past six months.

Childhood Affective Dysregulation Scale (CADS)[4,5]: Parent-report instrument designed to assess symptoms in youth with disruptive behavioral disorders. Includes the following subscales: aggression, impulsivity, mood lability, anger and irritability. Symptoms assessed over the past week.

Children's Hostility Inventory (CHI)[6]: Instrument designed to assess aggression and hostility in youth, and shown to distinguish conduct and non-conduct disorder diagnoses. A priori subscales include: Assaultiveness, Indirect Hostility, Irritability, Negativism, Resentment, Suspicion, and Verbal Hostility. Parent- and child-report; no time period is specified.

Disruptive Behavioral Disorders Rating Scale (DBD)[7]: Parent-report (and teacher-report) scale designed to assess for ADHD, ODD, and Conduct Disorder. Includes the following subscales: Opposition/Defiance, Inattention, and Hyperactivity. Symptoms assessed over the past four weeks.

Mood and Feelings Questionnaire (MFQ)[8,9]: Developed to assess depressive symptoms (including both cognitive and vegetative) in children and adolescents. Parent- and child-report; symptoms assessed over the past two weeks.

Screen for Child Anxiety Related Disorders (SCARED)[10]: Developed to assess anxiety symptoms in youth. We used subscales identified by a previous factor analysis: panic/somatic, generalized anxiety, separation anxiety, social phobia, and school phobia. Parent- and child-report; symptoms assessed over the past three months.

References

1. Birmaher B, Goldstein BI, Axelson DA, et al. Mood lability among offspring of parents with bipolar disorder and community controls. Bipolar Disord. May 2013;15(3):253-263.
2. Gerson AC, Gerring JP, Freund L, et al. The Children's Affective Lability Scale: a psychometric evaluation of reliability. Psychiatry Res. Dec 20 1996;65(3):189-198.
3. Achenbach TM. Manual for the Child Behavior Checklist/4-18 and 1991 profile. Burlington, VT: University of Vermont Department of Psychiatry; 1991.
4. Goldstein TR, Obreja M, Shamseddeen W, et al. Risk for suicidal ideation among the offspring of bipolar parents: results from the Bipolar Offspring Study (BIOS). Arch Suicide Res. 2011;15(3):207-222.
5. Kolko D, Baumann B, Chisholm C, E. B. The children's affective dysregulation scale: Initial psychometric properties in children with disruptive behavior disorders. Unpublished Manuscript; 2001.
6. Kazdin A, Rodgers A, Colbus D, Siegel T. Children's Hostility Inventory: Measurement of Aggression and Hostility in Psychiatric Inpatient Children. Journal of Clinical Child Psychology. 1987;16(4):320-328.
7. Pelham WE, Jr., Gnagy EM, Greenslade KE, Milich R. Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. J Am Acad Child Adolesc Psychiatry. Mar 1992;31(2):210-218.
8. Costello EJ, Angold A. Scales to assess child and adolescent depression: checklists, screens, and nets. J Am Acad Child Adolesc Psychiatry. Nov 1988;27(6):726-737.
9. Daviss WB, Birmaher B, Melhem NA, Axelson DA, Michaels SM, Brent DA. Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. J Child Psychol Psychiatry. Sep 2006;47(9):927-934.
10. Birmaher B, Brent DA, Chiappetta L, Bridge J, Monga S, Baugher M. Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. J Am Acad Child Adolesc Psychiatry. Oct 1999;38(10):1230-1236.

Criteria for Bipolar Disorder Not Otherwise Specified From the Course and Outcome of Bipolar Youth (COBY) Study ([1,2] adapted from [3])

    A. Child does not meet the *DSM-IV* criteria for bipolar-I or bipolar-II

    B. A distinct period of abnormally elevated, expansive, or irritable mood plus the following:

        1. At least 2 *DSM-IV-TR* "B" manic symptoms (3 if the mood is irritability only) that are clearly associated with the onset of abnormal mood

        2. A clear change in functioning

        3. the presence of the elated and/or irritable mood and manic symptoms for a significant part of the day (a *minimum* of 4 hours, though this did not necessarily need to be expressed consecutively)

        4. a *minimum* of 4 days (not necessarily consecutive) meeting criteria B.1-B.3 over the subject's lifetime

    C. Mood and affective symptoms must be abnormal for the child's level of development and environment

    D. Symptoms or mood changes that occur during substance use or antidepressant treatment do not count toward a bipolar diagnosis.

    E. Exclusion criteria:

        1. Current or lifetime *DSM-IV* diagnosis of schizophrenia, mental retardation, autism, or severe autism spectrum disorders

        2. Mood disorders due to substance abuse, a medical condition, or secondary to use of medications (e.g. corticosteroids).

    F. Subjects determined to have the onset of bipolar disorder before comorbid substance use disorders are included

    G. Subjects with mild comorbid Asperger disorder or pervasive developmental disorder–not otherwise specified are included if their mood symptomatology was clearly episodic and best accounted for by the bipolar diagnosis

Recently, Towbin et al.[3] proposed more restrictive criteria for Bipolar Disorder Not Otherwise Specified, which would limit the number of false positives in clinical practice. These were:

1. Recurrent (minimum of 4) distinct episodes meeting full criteria for a Manic or Hypomanic Episode, except for the duration criteria. Each episode must last at least one day, and at least one

episode must last a minimum of 2 consecutive days. For a day to "count" toward an episode, symptoms must be present for most of that day.

2. A Hypomanic Episode without a history of a Major Depressive Episode

These criteria were proposed long after the inception of BIOS, and thus the data from this study were not designed to assess these criteria directly. Specifically, we did not have data on duration of the longest episode; as a proxy for this, we have used maximum number of symptomatic days in a week.

Of the 55 youth with Bipolar Disorder Not Otherwise Specified at some point in the study (at baseline or follow-up), approximately 61% meet the more restrictive criteria. The rest did not meet the criteria because they had insufficient number of symptoms (11%) or they never had more than one day with manic symptoms during a single week (28%) (though they had multiple lifetime episodes). The children who met the new, more restrictive criteria were slightly older than those who did not (12 vs. 10.1 years old, p=.02); however, other than this, groups did not differ on demographics, functioning, manic/depressive symptoms, or conversion to bipolar-I/II. These results indicate that that these new, more restrictive criteria are very similar to the original criteria, at least in this at-risk sample.

We also note that over 98% of participants with Bipolar Disorder Not Otherwise Specified had experienced more than 7 cumulative lifetime days of Bipolar Disorder Not Otherwise Specified symptoms, and 84% had experienced more than 30 days. This indicates that the vast majority of our sample reported recurrent mood symptoms. Thus, it is not surprising that there were no differences in outcome between children who met the more restrictive Bipolar Disorder Not Otherwise Specified criteria and those who did not.

References

1.   Birmaher B, Axelson D, Strober M, et al. Clinical course of children and adolescents with bipolar spectrum disorders. Arch Gen Psychiatry. Feb 2006;63(2):175-183.
2.   Axelson D, Birmaher B, Strober M, et al. Phenomenology of children and adolescents with bipolar spectrum disorders. Arch Gen Psychiatry. Oct 2006;63(10):1139-1148.
3.   Towbin K, Axelson D, Leibenluft E, Birmaher B. Differentiating bipolar disorder-not otherwise specified and severe mood dysregulation. J Am Acad Child Adolesc Psychiatry. May 2013;52(5):466-481.

Variables Included in Factor Analysis and Models

Factor Analysis Imputation: To impute missing data, we used (1) all 82 variables in our three factor analyses + Mania Rating Scale; (2) demographics including baseline age, socioeconomic status, gender, race, living status, father's age at offspring's birth, mother's age at offspring's birth, and bipolar parent(s) vs. control parents; and (3) ofspring diagnoses including attention deficit hyperactivity disorder, depression, panic disorder, generalized anxiety disorder, social anxiety disorder, obsessive compulsive disorder, post-traumatic stress disorder, conduct disorder, oppositional defiant disorder, substance use disorder, and psychosis.

Cox and Logistic Regression Models: We adjusted for (1) demographics including baseline age, socioeconomic status, gender, race, living status, father's age at offspring's birth, mother's age at offspring's birth; (2) parental anxiety, depression, attention deficit hyperactivity disorder, disruptive behavioral disorder, and substance abuse disorders; and (3) offspring anxiety, depression, attention deficit hyperactivity disorder, disruptive behavioral disorder, and substance abuse disorders.

Lasso (Least Absolute Shrinkage and Selection Operator) Regression

For factor selection in Cox and logistic regressions, we used Lasso (programmed using the *penalized* package in R), which is a modified form of least squares regression that minimizes the sum of squared errors while constraining the sum of the absolute value of the coefficients to be less than a constant [Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. J. Royal. Statist. Soc B., Vol. 58, No. 1, pages 267-288]. In doing so, Lasso penalizes the absolute magnitude of regression coefficients to the extent some parameter estimates equate to zero. Thus, a large set of potentially correlated predictors can be reduced to smaller subset of predictors with nonzero parameter estimates. The amount of parameter estimate shrinkage is determined by a tuning parameter λ, whose value we selected using the L_1-regularization path algorithm published by Park and Hastie [Park, M. Y. and T. Hastie (2007). L_1-regularized path algorithm for generalized linear models. Journal of the Royal Statistical Society, Series B 69 (4), 659-677].

TABLE S1. Parent-Report Maximum Likelihood Factor Analysis. Items loading onto one factor (p>.3) are in yellow. Items loading onto more than one factor are in blue (factor chosen based on clinical consensus) and grey.

Scale	Subscale	Factor 1: Internalizing	Factor 2: Externalizing	Factor 3: Inattention/ Disinhibition
MFQ-P	Total Score	**0.49**	**0.51**	-0.04
CALS-P	Irritability	0.16	**0.71**	0.00
CALS-P	Mania	0.18	0.09	**0.44**
CADS	Aggression	-0.13	**0.95**	0.01
CADS	Irritability	0.09	**0.84**	-0.13
CADS	Anger	-0.02	**0.95**	-0.11
CADS	Lability	0.07	**0.87**	-0.16
CADS	Impulsivity	0.17	**0.56**	0.21
DBD	Opposition/Defiance	-0.05	**0.76**	0.17
DBD	Inattention/Hyperactivity	0.16	0.17	**0.51**
DBD	Crime	-0.13	**0.31**	0.15
SCARED-P	Somatic/panic	**0.77**	0.01	-0.08
SCARED-P	Anxiety	**0.83**	0.07	-0.10
SCARED-P	SAD	**0.55**	0.05	0.07
SCARED-P	Social phobia	**0.67**	-0.02	-0.09
SCARED-P	School phobia	**0.62**	0.07	0.03
CHI-P	Assaultiveness	0.00	**0.54**	0.11
CHI-P	Indirect Hostility	0.01	**0.39**	0.03
CHI-P	Irritability	0.15	**0.52**	-0.13
CHI-P	Negativism	0.05	**0.45**	0.17
CHI-P	Suspicion	0.26	**0.33**	0.01
CHI-P	Verbal Hostility	0.20	**0.40**	-0.03
CBCL	Anxious/Depressed	**0.83**	0.03	0.10
CBCL	Withdrawn	**0.62**	0.09	0.19
CBCL	Somatic Complaints	**0.63**	-0.02	0.16
CBCL	Social Problems	**0.56**	0.07	0.41
CBCL	Thought Problems	**0.55**	0.04	0.37
CBCL	Attention Problems	**0.30**	0.14	**0.55**
CBCL	Rule-Breaking Behav	0.01	**0.49**	**0.44**
CBCL	Aggressive Behavior	0.07	**0.71**	0.28

Child Affective Lability Scale (CALS), Mood and Feelings Questionnaire (MFQ), Scale for Child Anxiety Related Disorders (SCARED), Child Behavioral Checklist (CBCL), Disruptive Behavior Disorders (DBD) rating scale, Childhood Affective Dysregulation Scale (CADS), Children's Hostility Inventory (CHI)

TABLE S2. Child-Report Factor Analysis. Items loading onto one factor (p>.3) are in yellow. Items loading onto more than one factor are in blue (factor chosen based on clinical consensus) and grey.

Scale	Subscale	Factor 1: Internalizing	Factor 2: Externalizing	Factor 3: Affective Lability
MFQ-C	Total Score	**0.30**	**0.37**	**0.31**
CALS-C	Irritability	-0.08	0.22	**0.53**
CALS-C	Mania	-0.01	-0.02	**0.70**
CALS-C	Anxiety/Depression	0.14	-0.07	**0.73**
SCARED-C	Somatic/ Panic Total	**0.72**	-0.04	0.19
SCARED-C	Anxiety total	**0.56**	0.14	0.19
SCARED-C	Separation Anxiety Total	**0.72**	-0.02	0.02
SCARED-C	Social phobia Total	**0.69**	0.09	-0.11
SCARED-C	School phobia Total	**0.39**	0.16	0.26
CHI-C	Assaultiveness	0.09	**0.49**	0.10
CHI-C	Indirect hostility	-0.09	**0.52**	-0.05
CHI-C	Irritability	0.20	**0.45**	0.02
CHI-C	Negativism	-0.05	**0.48**	0.03
CHI-C	Resentment	0.20	**0.45**	0.02
CHI-C	Suspicion	0.10	**0.47**	0.11
CHI-C	Verbal hostility	0.15	**0.45**	0.01

Child Affective Lability Scale (CALS), Mood and Feelings Questionnaire (MFQ), Scale for Child Anxiety Related Disorders (SCARED), Children's Hostility Inventory (CHI)

TABLE S3. Depression Rating Scale Factor Analysis. Items loading onto one factor (p>.35) are in yellow. Items loading onto more than one factor are in blue (factor chosen based on clinical consensus) and grey.

Depression Rating Scale Item	Factor 1: Depression/ Atypical Symptoms	Factor 2: Sleep	Factor 3: Suicidality
Depressed Mood	0.84	0.02	0.06
Irritability, Anger	0.67	0.07	0.05
Reactivity of Dep or Irrit Mood	0.64	0.01	0.09
Diurnal Mood Var (AM Worse)	0.39	0.01	-0.22
Diurnal Mood Var (PM Worse)	0.39	0.12	-0.11
Excessive/Innapropriate Guilt	0.63	-0.07	0.07
Negative Self-Image	0.68	-0.01	0.18
Hopelessness, Pessimism	0.72	0.03	0.12
Aches and Pains	0.51	0.04	-0.06
Anhedonia, Low Motivation	0.78	0.09	0.03
Fatigue	0.73	0.11	-0.09
Inattention	0.38	0.16	0.04
Psychomotor Retardation	0.65	0.07	0.09
Social Withdrawal	0.77	0.01	0.07
Insomnia	-0.02	0.97	0.10
Initial Insomnia	-0.17	0.91	0.14
Middle Insomnia	0.03	0.43	-0.04
Terminal Insomnia	0.11	0.36	-0.01
Circadian Reversal	-0.01	0.35	-0.05
Non-Restorative Sleep	0.32	0.46	-0.09
Daytime Sleepiness	0.40	0.29	-0.05
Hypersomnia	0.65	-0.03	-0.18
Anorexia	0.37	0.08	0.25
Increased Appetite	0.46	-0.08	0.00
Craving for Sweets	0.33	0.01	-0.02
Weight Gain	0.33	-0.08	0.05
Leaden Paralysis	0.53	-0.06	0.08
Rejection Sensitivity	0.62	0.04	-0.01
Suicidal Ideation	0.44	0.11	0.41
Number of Suicidal Acts	0.13	0.00	0.50
Seriousness of Suicidal Acts	0.08	-0.07	0.84
Lethality of Suicidal Acts	-0.06	0.01	0.80
Recurrent Thoughts of Death	0.41	0.13	0.32

TABLE S4. Parent Demographic and Clinical Characteristics Across Group

	1. Bipolar Parents with Bipolar Offspring (n=31)	2. Bipolar Parents without Bipolar Offspring (n=188)	3. Community Controls (n=127)	Statistic	p	Bonferroni Adjusted Pairwise p-values		
						1 vs. 2	1 vs. 3	2 vs. 3
Demographics								
Male n (%)	1 (3.2)	39 (20.7)	29 (22.8)	$\chi 2$=6.17	**0.05**	0.1	0.1	~1
White n (%)	24 (77.4)	171 (91.0)	97 (76.4)	$\chi 2$=13.49	**0.001**	0.1	~1	**0.002**
Mean Age (SD)	36.5 (5.9)	39.5 (7.5)	41.0 (7.1)	F=5.19	**0.006**	0.09	**0.006**	0.2
Mean SES (SD)	30.7 (12.0)	34.7 (14.4)	37.4 (13.0)	F=3.44	**0.03**	0.4	**0.04**	0.3
Mean Offspring in Study (SD)	2.3 (1.1)	2.0 (1.0)	2.1 (0.9)	F=2.49	0.08			
Married n (%)	9 (29.0)	99 (52.7)	86 (67.7)	$\chi 2$=17.08	**0.0002**	0.05	**0.0007**	**0.03**
Proband Parent Diagnoses								
Bipolar-I n (%)	23 (74.2)	136 (72.3)	---	$\chi 2$=0.05	0.8			
Bipolar-II n (%)	8 (25.8)	52 (27.7)	---					
Any Depression n (%)	28 (90.3)	176 (93.6)	56 (44.1)	$\chi 2$=103.7	**<0.0001**	~1	**0.0003**	**<0.0001**
ADHD n (%)	9 (29.0)	49 (26.1)	3 (2.4)	$\chi 2$=32.37	**<0.0001**	~1	**0.0002**	**<0.0001**
Any Anxiety Disorder n (%)	26 (83.9)	143 (76.1)	34 (26.8)	$\chi 2$=84.87	**<0.0001**	~1	**<0.0001**	**<0.0001**
Disruptive Behavioral Disorder n (%)	14 (45.2)	70 (37.2)	7 (5.5)	$\chi 2$=45.60	**<0.0001**	~1	**<0.0001**	**<0.0001**
Substance Use Disorder n (%)	19 (61.3)	125 (66.5)	35 (27.6)	$\chi 2$=47.25	**<0.0001**	~1	**0.002**	**<0.0001**

TABLE S5. Coparental Diagnoses

	1. Offspring with Bipolar Spectrum (n=31)	2. At-Risk Offspring (n=188)	3. Community Control Offspring (n=127)	Statistic	p	Bonferroni Adjusted Pairwise p-values		
						1 vs. 2	1 vs. 3	2 vs. 3
Bipolar Spectrum Disorder n (%)	4 (12.1)	10 (3.1)	---	F=2.74	0.1			
Any Depression n (%)	6 (18.2)	85 (26.1)	25 (13.3)	F=4.09	**0.02**	~1	0.8	**0.01**
ADHD n (%)	1 (3.0)	9 (2.8)	4 (2.1)	F=0.11	0.9			
Any Anxiety Disorder n (%)	2 (6.1)	44 (13.5)	24 (12.8)	F=0.61	0.5			
Disruptive Behavioral Disorder n (%)	10 (3.0)	20 (6.1)	5 (2.7)	F=1.06	0.3			
Substance Use Disorder n (%)	10 (30.3)	105 (32.2)	38 (20.2)	F=2.57	0.08			

TABLE S6. Offspring Demographic Characteristics Across Group for (a) All Offspring and (b) Those With Follow-Up

(a)	1. Offspring with Bipolar Spectrum (n=33)	2. At-Risk Offspring (n=326)	3. Community Control Offspring (n=220)	Statistic	p	Bonferroni Adjusted Pairwise p-values		
						1 vs. 2	1 vs. 3	2 vs. 3
Male n (%)	13 (39.4)	163 (50.0)	99 (45.0)	1.11	0.3			
White n (%)	24 (72.7)	270 (82.8)	167 (75.9)	2.05	0.1			
Lives with Both Biological Parents n (%)	5 (15.2)	146 (44.8)	140 (63.6)	9.53	**<0.0001**	**0.03**	**0.0004**	**0.005**
Mean Mother's Age of Offspring's Birth (SD)	23.3 (4.8)	27.0 (5.8)	28.3 (5.7)	8.87	**0.0002**	**0.001**	**0.002**	0.2
Mean Father's Age of Offspring's Birth (SD)	26.8 (6.4)	29.7 (7.4)	30.5 (7.0)	2.63	0.07			
Mean Age at Intake (SD)	12.7 (2.7)	11.6 (3.6)	11.7 (3.4)	3.05	**0.05**	**0.04**	0.1	~1
Mean Age at Final Assessment (SD)	20.0 (4.2)	19.6 (4.8)	19.7 (4.1)	1.98	0.1			
Had at least 1 Follow-Up Assessment n (%)	30 (90.9)	310 (95.1)	213 (96.8)	1.52	0.2			
Child Diagnoses								
Bipolar-I n (%)	6 (18.2)	---	---					
Bipolar-II n (%)	5 (15.2)	---	---					
Bipolar, not-otherwise-specified n (%)	22 (66.7)	---	---					
Any Depression n (%)	15 (45.5)	67 (20.6)	13 (5.9)	17.45	**<0.0001**	**0.006**	**<0.0001**	**<0.0001**
ADHD n (%)	14 (42.4)	72 (22.1)	31 (14.1)	7.40	**0.0007**	**0.03**	**0.0006**	0.06
Any Anxiety Disorder n (%)	18 (54.6)	78 (23.9)	24 (10.9)	16.37	**<0.0001**	**0.001**	**<0.0001**	**0.0006**
Disruptive Behavioral Disorders n (%)	17 (51.5)	51 (15.6)	16 (7.3)	18.03	**<0.0001**	**<0.0001**	**<0.0001**	**0.01**
Substance Use Disorder n (%)	2 (6.1)	10 (3.1)	4 (1.8)	1.02	0.4			

(b)	1. Offspring with Bipolar Spectrum (n=30)	2. At-Risk Offspring (n=310)	3. Community Control Offspring (n=213)	F Statistic	p-value	Bonferroni Adjusted Pairwise Comparisons		
						1 vs. 2	1 vs. 3	2 vs. 3
Number of Follow-up Assessments (SD)	3.3 (1.3)	3.6 (1.3)	3.7 (1.1)	0.77	0.5			
Mean Duration of Follow-Up in Years (SD)	8.1 (2.0)	8.4 (2.5)	8.3 (2.3)	0.31	0.7			
Duration between Assessments (SD)	2.9 (1.6)	2.4 (0.9)	2.4 (0.8)	4.34	0.01	**0.03**	**0.01**	0.98

TABLE S7. Factor Scores Across Groups at Baseline, Adjusted for Demographics

		1. Offspring with Bipolar Spectrum (n=33)		2. At-Risk Offspring (n=326)		3. Community Control Offspring (n=220)		Type 3 Output		Bonferroni Adjusted Pairwise p-values		
		Mean*	SE	Mean*	SE	Mean*	SE	F Statistic	p-value	1 vs. 2	1 vs. 3	2 vs. 3
Parent Reports	Internalizing	5.78	0.85	0.96	0.32	-2.04	0.40	38.24	<0.0001	<0.0001	<0.0001	<0.0001
	Externalizing	7.69	1.16	0.76	0.40	-2.36	0.50	32.54	<0.0001	<0.0001	<0.0001	<0.0001
	Inattention/Disinhibition	1.25	0.25	0.22	0.08	-0.37	0.10	21.92	<0.0001	0.0003	<0.0001	<0.0001
Child Reports	Internalizing	2.12	0.44	0.19	0.15	-0.60	0.19	17.13	<0.0001	<0.0001	<0.0001	0.004
	Externalizing	1.36	0.35	0.08	0.12	-0.42	0.15	11.21	<0.0001	0.002	<0.0001	0.03
	Affective Lability	2.31	0.28	0.18	0.09	-0.36	0.12	37.94	<0.0001	<0.0001	<0.0001	0.001
Depression Rating Scale	Depressive/Atypical Symptoms	12.68	1.23	0.49	0.41	-2.64	0.51	66.12	<0.0001	<0.0001	<0.0001	<0.0001
	Sleep Problems	2.40	0.44	-0.066	0.15	-0.30	0.18	16.66	<0.0001	<0.0001	<0.0001	~1
	Suicidality	1.17	0.38	0.17	0.12	-0.46	0.15	10.78	<0.0001	0.04	0.0002	0.003
Mania Rating Scale	Total Standardized Scores	2.37	0.13	-0.04	0.04	-0.33	0.05	177.29	<0.0001	<0.0001	<0.0001	<0.0001

*Least Squares Means

TABLE S8. Factor Scores Across Groups at Baseline, Adjusted for Demographics and Parental Non-Bipolar Categorical Diagnoses

		1. Offspring with Bipolar Spectrum (n=33)		2. At-Risk Offspring (n=326)		3. Community Control Offspring (n=220)		Type 3 Output		Bonferroni Adjusted Pairwise p-values		
		Mean*	SE	Mean*	SE	Mean*	SE	F Statistic	p-value	1 vs. 2	1 vs. 3	2 vs. 3
Parent Reports	Internalizing	5.28	0.85	0.57	0.33	-1.42	0.43	23.45	**<0.0001**	**<0.0001**	**<0.0001**	**0.002**
	Externalizing	8.11	1.16	0.84	0.40	-2.54	0.51	37.61	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
	Inattention/Disinhibition	1.17	0.25	0.18	0.08	-0.28	0.10	15.38	**<0.0001**	**0.0004**	**<0.0001**	**0.003**
Child Reports	Internalizing	1.98	0.44	0.01	0.16	-0.33	0.20	11.33	**<0.0001**	**<0.0001**	**<0.0001**	0.6
	Externalizing	1.36	0.35	0.08	0.12	-0.42	0.15	11.21	**<0.0001**	**0.002**	**<0.0001**	**0.03**
	Affective Lability	2.35	0.28	0.13	0.10	-0.30	0.12	37.83	**<0.0001**	**<0.0001**	**<0.0001**	**0.02**
DRS	Depressive/Atypical Symptoms	12.74	1.23	0.14	0.42	-2.20	0.54	60.21	**<0.0001**	**<0.0001**	**<0.0001**	**0.004**
	Sleep Problems	2.40	0.44	-0.066	0.15	-0.30	0.18	16.66	**<0.0001**	**<0.0001**	**<0.0001**	~1
	Suicidality	1.17	0.38	0.17	0.12	-0.46	0.15	10.78	**<0.0001**	**0.04**	**0.0002**	**0.003**
MRS	Total Standardized Scores	2.36	0.13	-0.07	0.04	-0.28	0.06	170.96	**<0.0001**	**<0.0001**	**<0.0001**	**0.02**

*Least Squares Means

TABLE S9. Changes in Manic and Depressive Symptoms, As Well As Global Assessment of Function, Prior To, At, and After Conversion to Bipolar Spectrum Disorder

	1. Assessment Prior to Conversion (n=44)	2. Assessment at Conversion (n=44)	3. Assessment After Conversion (n=35)	F Stat	P-value	Tukey Adjusted Contrast P-values		
						1 vs. 2	1 vs. 3	2 vs. 3
Mean Depression Rating Score (SD)	7.3 (7.9)	11.3 (8.0)	12.1 (9.5)	4.31	**0.02**	**0.04**	**0.03**	0.95
Mean Mania Rating Score (SD)	6.3 (6.3)	15.6 (8.8)	13.6 (9.0)	23.66	**<0.0001**	**<0.0001**	**<0.0001**	0.2
Mean Global Assessment of Function (SD)	68.0 (13.7)	61.2 (10.1)	60.1 (10.6)	6.81	**0.002**	**0.006**	**0.006**	0.97

TABLE S10. Unpacking the *Parent-Report Internalizing* Score: Impact of individual scales on the hazard of developing bipolar spectrum within at-risk offspring, both separately and combined in one model.

		χ^2 Stat	p-value
Univariate Cox Models	Mood and Feelings Questionnaire	24.67	**<0.0001**
	Screen for Child Anxiety Related Disorders (all subscales)	27.02	**<0.0001**
	Child Behavioral Checklist (anxious/depressed, withdrawn, somatic complaints, social problems, thought problems)	35.07	**<0.0001**
Multiple Cox Model	Mood and Feelings Questionnaire	0.58	0.4
	Screen for Child Anxiety Related Disorders (all subscales)	0.18	0.7
	Child Behavioral Checklist (anxious/depressed, withdrawn, somatic complaints, social problems, thought problems)	5.81	**0.02**

TABLE S11. Unpacking the Baseline *Child-Report Affective Lability* Score: Impact of individual subscales on the hazard of developing bipolar spectrum within at-risk offspring, both separately and combined in one model.

		χ^2 Stat	p-value
Univariate Cox Models	Child Affective Lability Scale (Irritability)	13.28	**0.0003**
	Child Affective Lability Scale (Mania)	10.18	**0.001**
	Child Affective Lability Scale (Anxiety/Depression)	14.10	**0.0002**
Multiple Cox Model	Child Affective Lability Scale (Irritability)	4.40	**0.04**
	Child Affective Lability Scale (Mania)	0.70	0.4
	Child Affective Lability Scale (Anxiety/Depression)	2.32	0.1

TABLE S12. Unpacking the proximal *Child-Report Affective Lability* Score: Impact of individual subscales on the hazard of developing bipolar spectrum within at-risk offspring, both separately and combined in one model.

		χ^2 Stat	p-value
Univariate Cox Models	Child Affective Lability Scale (Irritability)	22.38	**<0.0001**
	Child Affective Lability Scale (Mania)	16.38	**<0.0001**
	Child Affective Lability Scale (Anxiety/Depression)	16.58	**<0.0001**
Multiple Cox Model	Child Affective Lability Scale (Irritability)	5.24	**0.02**
	Child Affective Lability Scale (Mania)	0.16	0.7
	Child Affective Lability Scale (Anxiety/Depression)	0.17	0.7

TABLE S13. Unpacking the *Mania Rating Scale* - Impact of individual items on the hazard of developing bipolar spectrum within offspring, both separately and combined in one penalized regression (Lasso) model.

		χ^2 Stat	p-value
Univariate Models	Elation	16.29	**<0.0001**
	Irritability, Anger	14.10	**0.0002**
	Decreased Need for Sleep	4.05	**0.04**
	Unusually Energetic	8.31	**0.004**
	Increased Goal-Directed Activity	6.63	**0.01**
	Motor Hyperactivity	24.01	**<0.0001**
	Grandiosity	12.56	**0.0004**
	Accelerated Speech	10.14	**0.001**
	Racing Thoughts	8.58	**0.003**
	Flight of Ideas	0.36	0.5
	Poor Judgment	7.77	**0.005**
	Inappropriate Laughing	12.40	**0.0004**
	People Seeking	1.23	0.3
	Increased Productivity	9.53	**0.002**
	Sharpened Creative Thinking	5.67	**0.02**
	Hypersexuality	1.81	0.2
	Distractibility	22.06	**<0.0001**
	Hallucinations	0.86	0.4
	Mood Lability	10.15	**0.001**
Selected Model	Irritability, Anger	4.72	**0.03**
	Motor Hyperactivity	6.24	**0.01**
	Distractibility	5.01	**0.03**

FIGURE S1. Symptoms of (a) depression, (b) mania, and (c) global function as measured at assessment prior to conversion, at conversion, and after conversion, shown separately according to bipolar disorder subtype (Bipolar-I/II vs. Bipolar not-otherwise-specified). No differences were significant between groups.
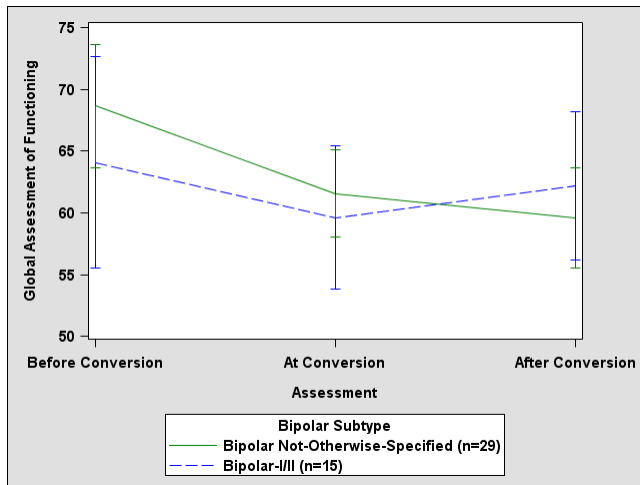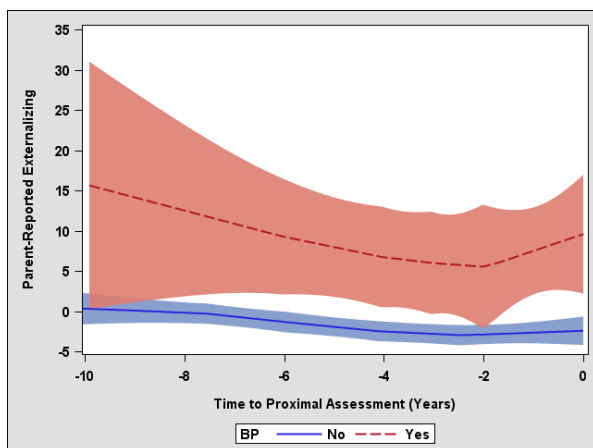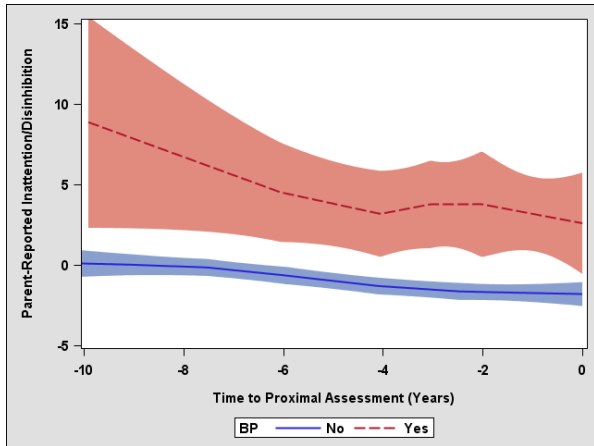
a.
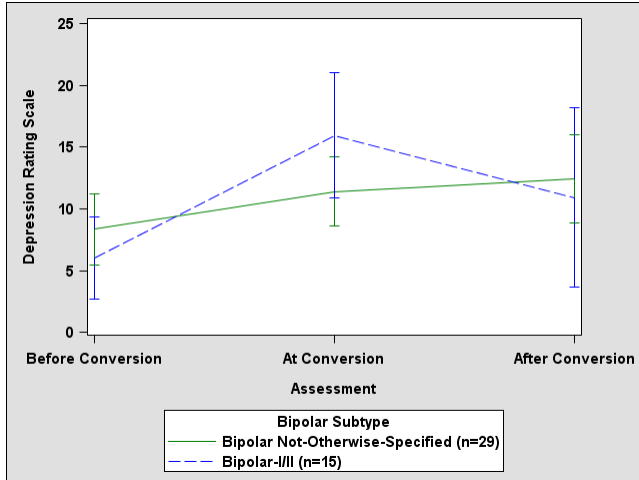
b.
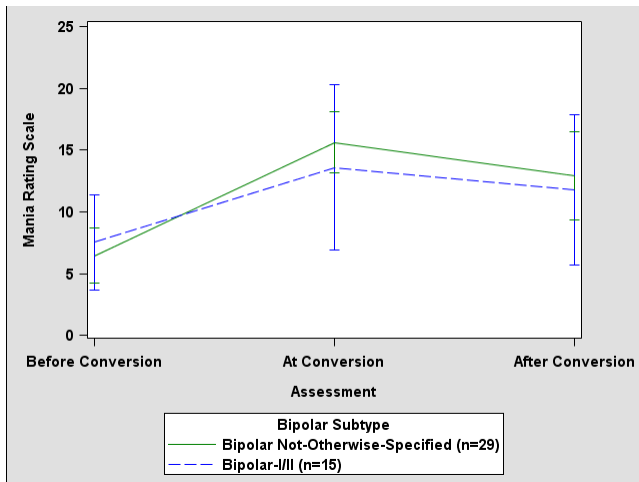
c.

FIGURE S2. Trajectories for dimensions graphed up to final eligible visit (either the visit prior to conversion or the visit prior to right censorship). a. Parent-reported externalizing, b. parent-reported inattention/disinhibition, c. child-reported internalizing, d. child-reported externalizing, e. clinically-derived depressive/atypical symptoms, f. clinically-derived sleep problems, and g. clinically-derived suicidality. While depressive/atypical symptoms and clinically derived suicidality appear to decrease up to the point of conversion, slopes were not significantly different from zero (and not significantly different across groups).
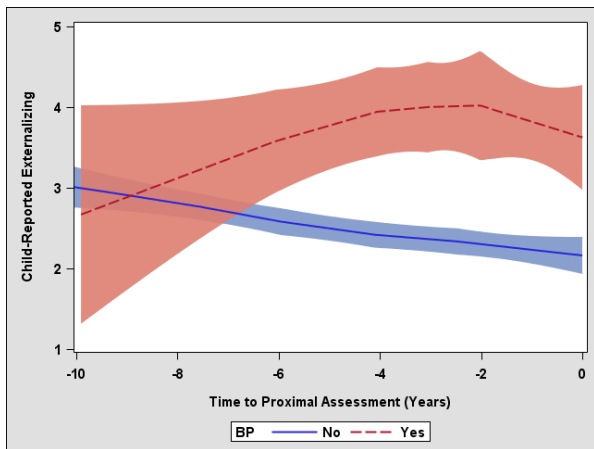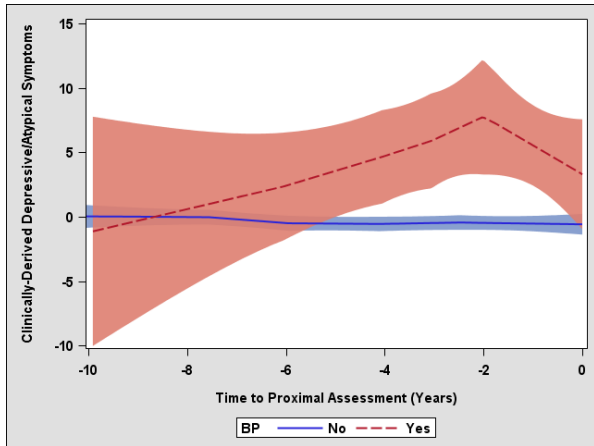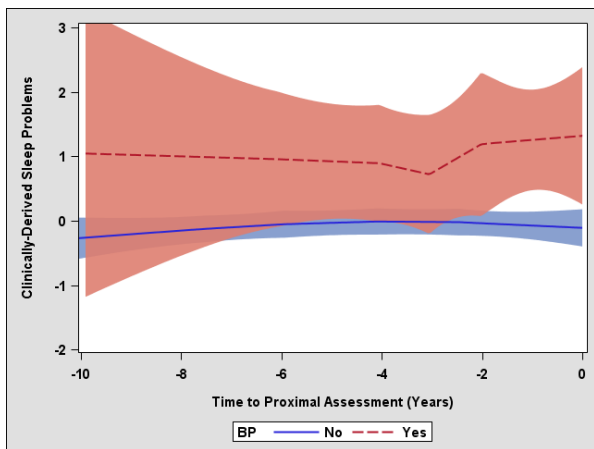
a.

b.

c.

d.

e.

f.

g.