



HHS Public Access

Author manuscript

J Am Acad Child Adolesc Psychiatry. Author manuscript; available in PMC 2021 October 01.

Published in final edited form as:

J Am Acad Child Adolesc Psychiatry. 2020 October ; 59(10): 1156–1164. doi:10.1016/j.jaac.2019.12.005.

Predicting Personalized Risk of Mood Recurrences in Youths and Young Adults With Bipolar Spectrum Disorder

Boris Birmaher, MD¹, John A. Merranko, MA¹, Mary Kay Gill, MSN¹, Danella Hafeman, MD¹, Tina Goldstein, PhD¹, Benjamin Goldstein, MD², Heather Hower, MSW³, Michael Strober, PhD⁴, David Axelson, MD⁵, Neal Ryan, MD¹, Shirley Yen, PhD³, Rasim Diler, MD¹, Satish Iyengar, PhD⁶, Michael W. Kattan, PhD⁷, Lauren Weinstock, PhD³, Martin Keller, MD³

¹Department of Psychiatry, Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania ²Department of Psychiatry, Sunnybrook Health Sciences Centre, University of Toronto Faculty of Medicine, Ontario, Canada ³Department of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University; Butler Hospital, Providence, Rhode Island ⁴Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, California ⁵Department of Psychiatry, Nationwide Children's Hospital and The Ohio State College of Medicine, Columbus, Ohio ⁶Department of Statistics, University of Pittsburgh, 2717 Cathedral of Learning, Pittsburgh, PA ⁷Cleveland Clinic Foundation, Cleveland, Ohio

Abstract

Objective: With each recurrence the prognosis of bipolar disorder (BD) worsens, indicating the need to identify the factors associated with increased recurrence risk. The course of BD is heterogenous and although risk factors for recurrence for the group as a whole have been reported in the literature, identification of risk factors for a specific individual are crucial for developing personalized treatments.

Method: 363 recovered BD youths/young adults from the Course and Outcome of Bipolar Youth (COBY) study were included. Participants were evaluated on average every 7-months for a median of 12.5 years and interviewed with standard instruments. Risk factors of recurrence from the

Correspondence to Boris Birmaher, MD, Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Bellefield Towers – Room 612, Pittsburgh, PA 15213; birmaherb@upmc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

This manuscript represents original material that has never been published before, is not under consideration for publication elsewhere, and has been approved by each author.

The work was completed at the Department of Psychiatry, Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania and the Department of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University.

This study was presented at Grand Rounds; April 10, 2019; the University of Cincinnati, Cincinnati, Ohio; Grand Rounds; April 19, 2019; the University of Wisconsin, Madison, Wisconsin; and at the World Congress of Biological Psychiatry 15th Biennial Conference; June 2-6, 2019; Vancouver, Canada.

literature were utilized to build a Risk Calculator (RC) to predict recurrence risk at different time intervals.

Results: Approximately 80% of participants had at least one syndromal recurrence and 60% had 2 recurrences, particularly depressions. The 6-month and 1, 2, 3, and 5-year RC showed an accuracy between 72% to 82% for predicting any mood recurrences, and up to 80% for depression and 89% for hypo/mania (sensitivity/specificity both 0.74). The most influential recurrence risk factors were shorter recovery lengths, younger age at assessment, earlier mood onset, and more severe prior depression. Although important, other factors associated with recurrence risk, such as interepisodic subsyndromal mood symptoms and comorbidities did not influence the RC score beyond factors noted above.

Conclusion: The RC provides a useful tool for predicting an individual's recurrence risk of depression and/or hypo/mania in BD youths and for developing personalized interventions and informing research. Replication studies are warranted.

Keywords

youth; bipolar disorder; longitudinal course; personalized medicine; risk calculator

Introduction

Bipolar Disorder (BD) is a recurrent illness that impacts psychosocial functioning and increases the risk for substance abuse and suicidality.^{1,2} Moreover, as each recurrence worsens prognosis, identifying persons at high risk of future episodes is a crucial need.

Among the factors associated with elevated risk of mood recurrences across age in BD are age, early age of mood onset, severity, number of past recurrences, subthreshold residual mood symptomatology, BD subtype, psychosis, comorbid disorders, lack of treatment and/or poor adherence, and family psychopathology (Supplement 1, available online).¹⁻¹⁴ However, since the course of BD is heterogeneous, with some people having more recurrences and others a more benign course, these predictors, while important at the *group level*, do not answer a crucial clinical question necessary to tailor treatments for a *specific* person: What is the risk of recurrence for an *individual* with BD?

Recently, person-specific mathematical models, referred to as risk calculators (RC), have been developed to identify the optimal group of factors that predict the likelihood that an individual will develop a specific condition in the future.^{15,16} For example, RCs have been successfully validated and implemented to predict the personalized risk of medical illness such as cancer or cardiovascular conditions.^{17,18} These RCs have become very useful tools to inform about the prognosis and treatment for a specific patient.¹⁵⁻¹⁸

Risk prediction models have been developed to predict the onset and persistence of depression, psychosis, suicidality, response to treatment for depression, medication side effects, and in one study, the onset of BD in depressed adult patients.^{15,16,19-25} In youths, the Pittsburgh Bipolar Offspring Study (BIOS) developed a RC to predict the 5-year risk of developing BD in offspring of parents with BD.²¹ The presence of subthreshold manic symptoms, depression, anxiety, and mood lability, poor functioning, and early parental age

of mood disorder onset predicted onset of BD with 76% discrimination. Finally, the Course and Outcome for Bipolar Youth (COBY) study, by including early-onset mood symptomatology, anxiety, mood lability, and family psychopathology, built and externally validated a RC which predicted the risk of developing BD I/II in youths with subthreshold manic symptoms to BD-I/II with 71% discrimination.²⁶ The above RC is useful for predicting new onset BD in children and adolescents who have not yet developed it and whose parents have BD. However, it cannot be used to predict recurrences in youths that already have this disorder or to predict the recurrence of depressive episodes.

In a prior publication, COBY reported the 4-year course and factors associated with increased risk for mood recurrences for the *group as a whole*.⁵ The goal of the current paper is, for the first time, to build a RC to predict the *personalized* risk of a depressive or hypo/manic recurrence for youths with BD who have been followed for over 12 years though the COBY study. To make the time-frame of this RC comparable to those published in medicine, we built our RC to predict the recurrence risk over 5-years following recovery of a mood episode, though other intervals of prediction were also analyzed.

Method

COBY's methods have been previously documented.⁴⁻⁶ Relevant to this paper, COBY initially enrolled 413 BD youths who fulfilled DSM-IV criteria for BD-I, and BD-II and an operationalized COBY criteria for BD-NOS (For the BD-NOS criteria, see Supplement 2, available online). To develop the RC, participants needed to have a history of a full-threshold mood episode and had to be in recovery (at least 2 months of no or minimal mood symptoms), which allowed 363 participants to be included in the RC sample (mean age at intake: 12.6 ± 3.2 ; mean age at the last follow-up: 24.5 ± 4.6 ; BD-I: $n=218$, BD-II: $n=25$, and BD-Not Otherwise Specified (NOS): $n=120$; Table 1). To be included in the RC analyses, participants with BD-NOS *at intake* were required to: 1) have had a full threshold mood episode, 2) have recovered from this episode for least 8 weeks, and 3) have at least one threshold mood episode after this recovery period.

Participants were interviewed approximately a median of every 7.4 months for a median duration of 12.5 years with an 88% rate of retention. As compared to retained participants, dropouts were less likely to live with both biological parents were more likely to have ADHD and family history of conduct disorder.

Participants were mainly recruited from outpatient clinics (67.6%) and directly interviewed for psychiatric disorders using the Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL).²⁷ Participants with schizophrenia, $IQ < 70$, autism, and mood disorders secondary to substances, medications, or medical conditions were excluded.

Week-by-week longitudinal change in psychiatric symptoms and treatment was assessed using the Longitudinal Interval Follow-up Evaluation (LIFE) and quantified using the instrument's Psychiatric Status Rating (PSR) scale.²⁸ The PSR uses numeric values linked to the DSM-IV²⁹ criteria and participants' functioning. For mood disorders, PSR=1-2 indicates

no or minimal symptoms, 3-4 subthreshold mood symptoms, and ≥ 5 for syndromal symptomatology (For details regarding the PSR scoring see Supplement 3, available online). As defined in the literature (e.g.,³⁰), recovery was defined as two consecutive months with PSR=1-2 and recurrence by a PSR score of ≥ 5 for at least one week for hypomania/mania/mixed symptoms, and at least two weeks for major depressive episode. The consensus scores obtained after interviewing parents and their children were used for the analyses.

Parents were evaluated using the Structured Clinical Interview (SCID),³¹ and were interviewed about first/second-degree family members with a modified Family History Screen.³² The family psychopathology presented in this paper represents the summary of data collected during the full length of the study.

Socio-economical-status was ascertained using the Hollingshead Scale.³³

Research staff were trained to reliably administer the interviews, and child psychiatrists/psychologists confirmed all diagnoses during consensus conferences. The overall KSADS-PL kappas for psychiatric disorders were ≥ 0.8 and for the subthreshold mood disorders ascertained through the PSR ≥ 0.75 .

Each participating University's Institutional Review Board approved the study, and consent and/or assent was obtained from the participants and their parents.

Statistical Analysis

Kaplan-Meier estimation and mixed regressions were used for descriptive analysis of recurrence rates and predicted risk scores.

The RC was built following the criteria recommended by Fusar-Poli and colleagues.¹⁶ Participants' recovery periods were divided into six-month intervals beginning with the onset of the recovery, which enabled analysis of the effect of current recovery length on future recurrence-risk. Each datapoint defined the multinomial outcome variable as recurrence status (depressive recurrence vs. hypo/manic recurrence vs. no recurrence) over the next 5-years from each distinct point during recovery.

To avoid overfitting the model, the RC was built with predictors based on factors from the existing literature, rather than the results from COBY or the results of the prior RC that predicted risk of developing BD I/II in youths with subthreshold manic symptoms.²⁶ To optimize the practicality and disseminability of the RC, the most reproducible risk factors published in the literature were included: past and present duration of the remissions (or the corresponding, number of recurrences), severity of mood symptoms (i.e., maximum PSR scores recorded during the previous episode), age, age of mood onset, and first and/or second degree family history of mania (Supplement 1, available online).¹⁻¹⁴ In addition, we also explored the influence on the RC of other factors that have been associated with increased risk for recurrence such as interepisodic subsyndromal mood symptoms, suicidality, psychosis, abuse, comorbidities (e.g., ADHD), and non-manic/hypomanic family psychopathology (Supplement 1, available online).

Initial prediction models implemented various survival analysis designs, such as competing risk regression³⁴ and recurrent events Cox regression.³⁵ However, these parametric techniques failed to account for both the multinomial and recurrent nature of the outcome variable. Thus, the final model was trained via boosted multinomial classification trees,^{36,37} a useful model for these data since it implicitly incorporates interactions between predictors and has been previously demonstrated to effectively predict depressive episodes.²⁵ Predictions were then calibrated via Platt scaling.³⁸ To avoid overfitting, approximately half of the sample was used to train the model, while the other half was used to independently test the model. Further, the training sample was divided into five folds so that Platt scaling could be performed via cross-validation.³⁹ An algorithm was employed to optimize balance between randomized sets and folds in number of observations and recurrences per subsample. Randomization was performed at the participant level, meaning that when a participant was randomized into a given fold, all of his/her observations were assigned to that one fold. For additional statistical information, see Supplement 4, available online.

To test the validity of the model, discrimination was assessed via the area under the receiver operating characteristic curve (AUC) on the testing sample (evaluating predicted probability of any recurrence). Test sensitivity, specificity, and positive predictive value were assessed at a range of risk-thresholds. Calibration was tested via Hosmer-Lemeshow testing (by predicted risk decile) and by plotting and comparing observed vs. predicted recurrence-risk. To test the predictive importance of each variable, two measures were used: (1) Friedman's relative influence statistic³⁷ (calculated via cross-validation) and (2) the decrement in AUC with removal of that variable from the model (significance of decrements tested via 95% bootstrap interval). After the final model was trained and tested, the addition of other factors that have been associated with increased risk for mood recurrences were evaluated (Supplement 1, available online).

The effectiveness of the RC to predict recurrences was also evaluated for 6 months and 1,2,3, and 5 years. Finally, sensitivity analyses were also performed to assess the RC including only BD-I/II observations, and then only BD-I observations, as well as considering PSR = 4 instead of PSR = 5, because a PSR= 4 includes participants with substantial subthreshold mood symptoms.

Results

Rates of Recovery/Recurrences.

During a median follow-up time of 12.5 years (mean age at the last assessment: 24.5±4.6), participants had a median of two recovery periods with a median recovery length of 2.1 years with a range between 8 weeks to 15 years. Over 80% (294/363) of participants had at least one syndromal recurrence (Kaplan-Meier estimated risk = 83%), and 204 experienced 2 recurrences (estimated risk = 60%; range: 2-15), with most recurrences being depressions (71%).

The estimated median time until first recurrence after initial recovery was 1.5 years (range: 56 days – 11.8 years). The proportion of participants who had depressive or hypo/manic recurrences increased over time, but at progressively decreasing rates, until approximately

the three-year point, after which cumulative incidence stabilized (Figure S1, available online). Thus, for the RC analyses, participants' six-month recovery periods were truncated at an upper bin of 3+ years.

Five-year Risk Calculator.

The training/testing subsample randomization resulted in a training set of 182 participants with 1155 observations and a testing set of 181 participants with 1263 observations. There were no between-subsample significant differences in length of follow-up, percent of follow-up in recovery, length of recovery periods, number of recurrences, polarity of recurrences, or on any demographic, clinical, or family psychiatric history factors. Overall, the training/testing set partitioning procedure yielded subsamples with a high degree of balance. The test AUCs of the boosted multinomial regression's risk predictions were 0.82 (95% bootstrap interval: 0.81, 0.84), 0.80 (0.78,0.82) and 0.89 (0.85,0.91) for any mood, depressive, and hypo/manic episodes, respectively. The calibration plot (Figure S2, available online) indicated that predicted and observed recurrence-risk were consistent through the range of risk-scores, and the median predicted 5-year risk in the testing sample (0.52) closely matched the observed rate of recurrences in the testing sample (event rate=0.55). Further, predicted risk and observed rates of conversion within decile did not significantly differ (Hosmer-Lemeshow $\chi^2=5.33$, $df=8$, $p=0.7$), indicating no evidence of miscalibration. Table 2 shows test prediction metrics at a range of predicted risk thresholds. Optimal sensitivity and specificity were both 0.74 at a predicted risk threshold of approximately 0.50, with positive predictive value of 0.78. Of the eleven factors included in the model (Table 3), five accounted for over 80% of the overall relative influence to increase the recurrence risk. In order of influence, the five factors were: 1) prior recovery lengths (higher risk with shorter recoveries), 2) age at the time of the assessment (higher risk with younger age), 3) current recovery length (higher risk with shorter recovery), 4) severity of the previous DSM major depressive episode (PSR = 5), and 5) age of mood disorder onset (higher risk with earlier onset). The only factors associated with a significant decrement in the AUC when removed were age at the time of the assessment, age of mood onset, and current recovery length.

When evaluating predictor importance differentially by polarity of the next recurrence, current recovery length was most predictive of depressive recurrences, presence of hypo/manic symptoms in most prior episodes was most predictive of manic episodes, and age, age of mood disorder onset, and prior recovery length were strong predictors of both depressive and manic recurrences.

In addition to the eleven factors noted above, we explored the influence of other factors that have been reported in the literature to be associated with increased recurrence-risk including history of suicide attempts, psychosis, inter-episodic subthreshold mood symptoms, polarity of the index episode, abuse, comorbid disorders (e.g., ADHD, SUD), and family history of unipolar depression and non-mood psychopathology (Table 4). These factors exerted minimal additional influence in the model above and beyond the effects of the eleven factors noted above (Table 3).

Risk Calculator for 6-months and 1, 2, and 3-years (Figure 1):

The RC using the same predictors in the 5-year model was also effective when predicting any mood recurrence-risk at 6 months and 1, 2, and 3 years after recovery with AUCs of 0.72, 0.75, 0.78, and 0.80, respectively, and showed strong calibration (Hosmer-Lemeshow p-values ranging from 0.3-0.7). Similar AUCs were found for depressive recurrences. For hypomanic/manic recurrences, the AUCs were around 0.89 for all time intervals.

Sensitivity Analyses.

AUCs were comparable when training and testing the model including only BD-I/II observations (0.83) and then only BD-I observations (0.85), and the same above-noted five predictors made up the top 80% of overall relative influence in both retrained models. Also, the model was trained and tested when the definition of recurrence included subthreshold manic symptoms (PSR 4) instead of threshold manic symptoms (PSR 5) yielding an AUC = 0.82.

Low Base Rate but Clinically Severe Risk Factors.

The relative influences of suicide attempts and psychosis on the RC were less than 1%, indicating low predictive influence relative to the other factors in the model (Table 4). However, given the clinical significance of these factors, we further analyzed whether adding suicide attempts or psychosis to the model would substantially change estimated risk scores in the subsample with these factors. Comparing standardized risk scores estimated by models with versus without each factor in the model amongst participants with the respective factor, we found that the mean difference in standardized risk-scores was virtually zero (both p-values~1). Further, linear mixed regressions found that participants with history of psychosis and suicide attempts already had significantly higher risk-scores, as estimated by the original RC (p-values<0.01).

Discussion

Similar to the youth and adult literature,^{1,7,9,11,12} we found that among recovered BD youth during a median of 12.5 years of follow-up, about 80% had at least one syndromal recurrence, and 60% experienced 2 recurrences, particularly major depressions. Participants had a median of two recovery periods during follow-up, and the median recovery length was 2.1 years. Using eleven predictors of mood recurrence from the BD literature that are commonly ascertained in clinical practice (Table 3), we successfully built a practical and potentially, disseminable RC. Utilizing these factors, the 5-year RC showed excellent discrimination with an AUC of approximately 0.82 for any recurrences and 0.8 to 0.89 for the depression and hypo/manic recurrences, respectively, and a sensitivity and specificity both at 0.74 with a positive predictive value of 0.78. Similar findings were found when the analyses were done with only BD-I youths or when the definition of recurrence included significant subthreshold mood symptoms (PSR 4) instead of threshold symptoms (PSR 5). The AUCs for any mood recurrence-risk at 6 months and 1,2, and 3 years ranged from 0.72 to 0.80 (Figure 1). Similar findings were observed for depressive recurrences. For hypo/manic recurrences, the AUCs were around 0.89 for all time intervals.

Of the eleven factors associated with increased risk recurrence included in the RC, five accounted for over 80% of the overall relative influence to increase the recurrence risk including shorter current and previous recovery lengths, younger age at the time of assessment, earlier mood onset, and more severe depression in the previous episode. When predicting differentially by episode polarity, current recovery length was most predictive of major depressive recurrences, and presence of hypo/manic symptoms in most prior episodes was most predictive of hypo/manic episodes. This last finding corroborated prior reports indicating that the polarity of the initial mood episodes predicts the polarity of future episodes.^{1,2,5}

It is important to emphasize that there are other factors that have been reported in the BD literature to be associated with recurrence risk (e.g., inter-episodic subsyndromal mood symptoms, comorbid disorders, trauma, suicidality, psychosis) (Supplement 1, available online).¹⁻¹⁴ Each of these factors in isolation are important, but perhaps due to the fact that they tend to co-occur with the factors included in the RC (e.g., more severe depressive episodes are associated with more suicidality), they did not influence the RC score *above and beyond* the factors included in the model. For example, the recurrence-risk was not underestimated for participants with suicidality or psychosis because even without including these risk factors in the model, these participants had significantly higher risk scores. However, it is important to note that given the severity of some of the symptoms (e.g., suicidality and psychosis), clinicians are advised to consider the gravity of these conditions when evaluating a patient with a low risk score.

The proposed RC showed AUCs that are comparable to RCs developed to predict psychosis, new onset of BD in offspring of parents with BD, and medical illnesses (AUCs ~ 0.7-0.8).¹⁵⁻²⁵ Thus, it can be used for biological and treatment studies, such as comparing the biological differences and differential treatment response between youths/young adults with low and high-risk scores. A clinician could utilize the RC, which already considers the inherent heterogeneity of the course of BD,^{1-3,9,11,12} to predict the *person-level risk* for recurrences by polarity at different points of time after recovery. This information can inform the management of individual patients and is valuable for the patients and their families to enhance their understanding of illness prognosis and become more involved in their treatment. For example, a high-risk score may indicate the need for more frequent follow-ups, and more intensive treatment to avoid recurrences.

The results of this study should be considered within the context of the following limitations. Most participants were Caucasian recruited from University settings limiting the generalizability of the results. Nonetheless, course and outcome in non-clinically referred BD youth have been shown to be similar to those in referred populations.⁴⁰ COBY recruited participants with childhood/adolescent-onset BD. Thus, it is unclear to what extent the results apply to adult-onset BD. Type of treatment and adherence to treatment can also influence the risk of recurrence.⁴¹ However, we did not measure adherence and since the effects of treatment are confounded by indication, treatment was not included in the RC. The monitoring of symptoms using the PSR in research settings, and in our study, is more systematic than what is typically done in clinical practice. Thus, the need to externally validate the calculator in routine clinical practices. Despite that COBY is a prospective

longitudinal study, the assessment of symptoms since the last assessment are retrospective and subject to recollection bias. Finally, the presence of factors associated with high-risk for recurrence are not stable and may change over time, and as new findings are available, new factors (e.g., genetic and neuroimaging) may be added.

In conclusion, although knowing the factors associated with recurrence-risk across a BD sample is informative, given the heterogeneity of the course of BD, determining the individual recurrence-risk is critical to personalize the treatment for a specific-person. For this purpose, we proposed a simple, cost-effective, and disseminable RC to predict mood recurrences and the polarity of the recurrences that includes a limited number of self-explanatory factors that usually are part of the regular management of BD. Thus, the RC offers a potentially useful clinical and research tool to predict the individual progression of the illness and guide treatment. Further validation in an independent sample is warranted.

For illustration, the RC instructions for obtaining information about mood PSR scores and other factors may be found at www.pediatricbipolar.pitt.edu (under Resources), and are also included in the Supplements 3 and 5 available online. The following vignettes are examples of the clinical use of the RC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This paper was supported by grants from the National Institute of Mental Health (NIMH): MH059929 (Birmaher), MH059691 (Keller/Yen), MH059977 (Strober), MH112544 (Birmaher), and MH112543 (Yen).

Mr. Merranko and Dr. Iyengar served as the statistical experts for this research.

The authors would like to thank the studies' participants and their families, the research assistants, and Rita Scholle, BA, of UPMC Western Psychiatric Hospital, for preparation of the manuscript. The authors would also like to acknowledge Stacia Friedman-Hill, PhD, and Shelli Avenevoli, PhD, of NIMH, for their continued encouragement and support.

Disclosure: Dr. Birmaher has reported grants from NIMH, during the conduct of the study, and royalties from Random House, UpToDate, and Lippincott, Williams, and Wilkins, outside of the submitted work. Dr. Hafeman has reported grants from NIMH and the Klingenstein Third Generation Foundation. Dr. T. Goldstein has reported grants from NIMH, the American Foundation for Suicide Prevention, the University of Pittsburgh Clinical and Translational Science Institute (CTSI), and the Brain and Behavior Foundation, and royalties from Guilford Press, outside the submitted work. Dr. B. Goldstein has received grant or research support from NIMH, the Brain and Behavior Research Foundation (NARSAD), Brain Canada, the Canadian Institutes of Health Research, the Heart and Stroke Foundation of Canada, the Ontario Ministry of Research and Innovation, and the University of Toronto. Dr. Strober has received research support from NIMH and support from the Resnick Endowed Chair in Eating Disorders. Dr. Axelson has received grants from NIMH during the conduct of the study, and personal fees from Janssen Research and Development, LLC and UpToDate, outside the submitted work. Dr. Ryan has received grant or research support from NIMH. He has served on the Scientific Advisory Board of the Child Mind Institute. Dr. Yen has received research support from NIMH and the American Foundation for Suicide Prevention and has served as a consultant at Janssen Research and Development, LLC. Dr. Diler has received research support from NIMH. Dr. Kattan has reported personal consulting for Exosome Diagnostics, Novartis, Glaxo Smith Kline, and Boehringer Ingelheim and has had research sponsored by Novo Nordisk, Otsuka, and Merck. Dr. Weinstock has received research support from NIMH. Dr. Keller has received research support from NIMH. Ms. Hower has received research support from NIMH and honoraria from the Department of Defense (DOD). Mr. Merranko, Ms. Gill, and Dr. Iyengar have reported no biomedical financial interests or potential conflicts of interest.

References

1. Birmaher B Bipolar Disorders In: Martin A, Volkmar FR, ed. *Lewis's Child and Adolescent Psychiatry: A Comprehensive Textbook*. 5th ed. Philadelphia: Wolters Kluwer; 2018.
2. Grande I, Berk M, Birmaher B, Vieta E. Bipolar disorder. *Lancet*. 2016;387(10027):1561–1572. [PubMed: 26388529]
3. Altman S, Haeri S, Cohen LJ, et al. Predictors of relapse in bipolar disorder: A review. *J Psychiatr Pract*. 2006;12(5):269–282. [PubMed: 16998415]
4. 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 e1003. [PubMed: 21961775]
5. Birmaher B, Axelson D, Goldstein B, et al. Four-year longitudinal course of children and adolescents with bipolar spectrum disorders: the Course and Outcome of Bipolar Youth (COBY) study. *Am J Psychiatry*. 2009;166(7):795–804. [PubMed: 19448190]
6. Birmaher B, Gill MK, Axelson DA, et al. Longitudinal trajectories and associated baseline predictors in youths with bipolar spectrum disorders. *Am J Psychiatry*. 2014;171(9):990–999. [PubMed: 24874203]
7. DelBello MP, Hanseman D, Adler CM, Fleck DE, Strakowski SM. Twelve-month outcome of adolescents with bipolar disorder following first hospitalization for a manic or mixed episode. *Am J Psychiatry*. 2007;164(4):582–590. [PubMed: 17403971]
8. Fiedorowicz JG, Endicott J, Solomon DA, Keller MB, Coryell WH. Course of illness following prospectively observed mania or hypomania in individuals presenting with unipolar depression. *Bipolar disorders*. 2012;14(6):664–671. [PubMed: 22816725]
9. Goldberg JF, Garno JL, Harrow M. Long-term remission and recovery in bipolar disorder: a review. *Current psychiatry reports*. 2005;7(6):456–461. [PubMed: 16318824]
10. Hong W, Zhang C, Xing MJ, et al. Contribution of long duration of undiagnosed bipolar disorder to high frequency of relapse: A naturalistic study in China. *Compr Psychiatry*. 2016;70:77–81. [PubMed: 27624425]
11. Kessing LV, Andersen PK, Vinberg M. Risk of recurrence after a single manic or mixed episode - a systematic review and meta-analysis. *Bipolar disorders*. 2018;20(1):9–17. [PubMed: 29239075]
12. Radau J, Grunze H, Amann BL. Meta-Analysis of the Risk of Subsequent Mood Episodes in Bipolar Disorder. *Psychother Psychosom*. 2017;86(2):90–98. [PubMed: 28183076]
13. Tundo A, Musetti L, Benedetti A, et al. Predictors of recurrence during long-term treatment of bipolar I and II disorders. A 4 year prospective naturalistic study. *J Affect Disord*. 2018;225:123–128. [PubMed: 28826088]
14. Vazquez GH, Holtzman JN, Lolich M, Ketter TA, Baldessarini RJ. Recurrence rates in bipolar disorder: Systematic comparison of long-term prospective, naturalistic studies versus randomized controlled trials. *Eur Neuropsychopharmacol*. 2015;25(10):1501–1512. [PubMed: 26238969]
15. Bernardini F, Attademo L, Cleary SD, et al. Risk prediction models in psychiatry: Toward a new frontier for the prevention of mental illnesses. *J Clin Psychiatry*. 2017;78(5):572–583. [PubMed: 27337225]
16. Fusar-Poli P, Hijazi Z, Stahl D, Steyerberg EW. The Science of Prognosis in Psychiatry: A Review. *JAMA psychiatry*. 2018;75(12):1289–1297. [PubMed: 30347013]
17. Ankerst DP, Hoefler J, Bock S, et al. Prostate Cancer Prevention Trial risk calculator 2.0 for the prediction of low- vs high-grade prostate cancer. *Urology*. 2014;83(6):1362–1367. [PubMed: 24862395]
18. Goff DC Jr., Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129(25 Suppl 2):S49–73. [PubMed: 24222018]
19. Cannon TD, Yu C, Addington J, et al. An Individualized Risk Calculator for Research in Prodromal Psychosis. *Am J Psychiatry*. 2016;173(10):980–988. [PubMed: 27363508]

20. Chekroud AM, Zotti RJ, Shehzad Z, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *The Lancet Psychiatry*. 2016;3(3):243–250. [PubMed: 26803397]
21. Hafeman DM, Merranko J, Goldstein TR, et al. Assessment of a person-level risk calculator to predict new-onset bipolar spectrum disorder in youth at familial risk. *JAMA psychiatry*. 2017;74(8):841–847. [PubMed: 28678992]
22. Klein NS, Holtman GA, Bockting CLH, Heymans MW, Burger H. Development and validation of a clinical prediction tool to estimate the individual risk of depressive relapse or recurrence in individuals with recurrent depression. *J Psychiatr Res*. 2018;104:1–7. [PubMed: 29908389]
23. Liu Y, Sareen J, Bolton JM, Wang JL. Development and validation of a risk prediction algorithm for the recurrence of suicidal ideation among general population with low mood. *J Affect Disord*. 2016;193:11–17. [PubMed: 26766029]
24. Melhem NM, Porta G, Oquendo MA, et al. Severity and Variability of Depression Symptoms Predicting Suicide Attempt in High-Risk Individuals. *JAMA psychiatry*. 2019.
25. Van Voorhees BW, Paunesku D, Gollan J, Kuwabara S, Reinecke M, Basu A. Predicting future risk of depressive episode in adolescents: the Chicago Adolescent Depression Risk Assessment (CADRA). *Ann Fam Med*. 2008;6(6):503–511. [PubMed: 19001302]
26. Birmaher B, Merranko JA, Goldstein TR, et al. A Risk Calculator to Predict the Individual Risk of Conversion From Subthreshold Bipolar Symptoms to Bipolar Disorder I or II in Youth. *J Am Acad Child Adolesc Psychiatry*. 2018;57(10):755–763 e754. [PubMed: 30274650]
27. 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.[see comment]. *J Am Acad Child Adolesc Psychiatry*. 1997;36(7):980–988. [PubMed: 9204677]
28. 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. [PubMed: 3579500]
29. American Psychiatric Association (APA). *Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)*. Washinton, DC: American Psychiatric Association; 2004.
30. 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. [PubMed: 12044195]
31. First M, Spitzer RL, Gibbon M, Williams JBW. *Structures Clinical Interview for DSM-IV Axis I Disorders, Clinician Version (SCID-CV)*. Washington, D.C.: American Psychiatric Press Inc; 1996.
32. Weissman MM, Wickramaratne P, Adams P, Wolk S, Verdelli H, Olfson M. Brief screening for family psychiatric history: the family history screen. *Arch Gen Psychiatry*. 2000;57(7):675–682. [PubMed: 10891038]
33. Hollingshead AB. *Four-Factor Index of Social Status*. New Haven, Connecticut: Yale University Department of Sociology; 1975.
34. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*. 1999;94(446):496–509.
35. Andersen PK, Gill RD. Cox's Regression Model for Counting Processes: A Large Sample Study. *Ann Statist*. 1982;10(4):1100–1120.
36. Freund Y, Schapire RE. A Short Introduction to Boosting. *Journal of Japanese Society for Artificial Intelligence*. 1999;14(5):771–780.
37. Friedman JH. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*. 2001;14(5):771–780.
38. Niculescu-Mizil A, Caruana R. Obtaining calibrated probabilities from boosting. *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*; 2005; Edinburgh, Scotland.
39. Platt JA. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods. In: *Classifiers AILM*, ed.1999.
40. Lewinsohn PM, Klein DN, Seeley JR. Bipolar disorder during adolescence and young adulthood in a community sample. *Bipolar disorders*. 2000;2(3 Pt 2):281–293. [PubMed: 11249806]

41. Goldstein TR, Krantz M, Merranko J, et al. Medication Adherence Among Adolescents with Bipolar Disorder. *J Child Adolesc Psychopharmacol*. 2016;26(10):864–872. [PubMed: 27419273]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Clinical Vignette

Adam is a 17-year-old male with history of bipolar disorder (BD) since he was 10 years old, and he has both a parent and a grandparent with BD. His last episode was 2 years ago and lasted about 12 weeks. This episode was manifested with depressive symptoms, and the most severe depressive score (using the rating scale provided in eSupplement 3 was 5. Before the last mood episode, Adam was well for one year. Since age 10, he has had more than 2 episodes mainly manifested by depression. Entering this information into the risk calculator yielded risk scores of 0.19, 0.39, 0.66, and 0.73 for years 1, 2, 3, and 5, respectively. These results indicated that Adam was at 19%, 39%, 66%, and 73% predicted risk to have a mood recurrence in the next 1, 2, 3, and 5 years, respectively. Overall, the 5-year predicted risk of a depressive episode is 69% as compared to 4% for a manic episode.

Based on the above information, Adam's therapy and pharmacological treatment should focus mainly on preventing depression. Adam and his family should be informed that he is at high risk to have a depressive episode. Thus, he and his family should be advised to seek treatment as soon as the depressive symptoms begin to occur and educated about the importance of adhering to treatment to avoid further recurrences.

In contrast, Jane is a 16-year-old patient with similar clinical characteristics to Adam, but has infrequent mood episodes and no family history of bipolar disorder. The risk calculator for Jane showed very low 1-5-year risk scores, and consequently, much less likelihood of having a recurrence in the next 5 years. This case raises the question of whether patients with very low risk for recurrence should be treated with ongoing pharmacotherapy, an issue that warrants further research.

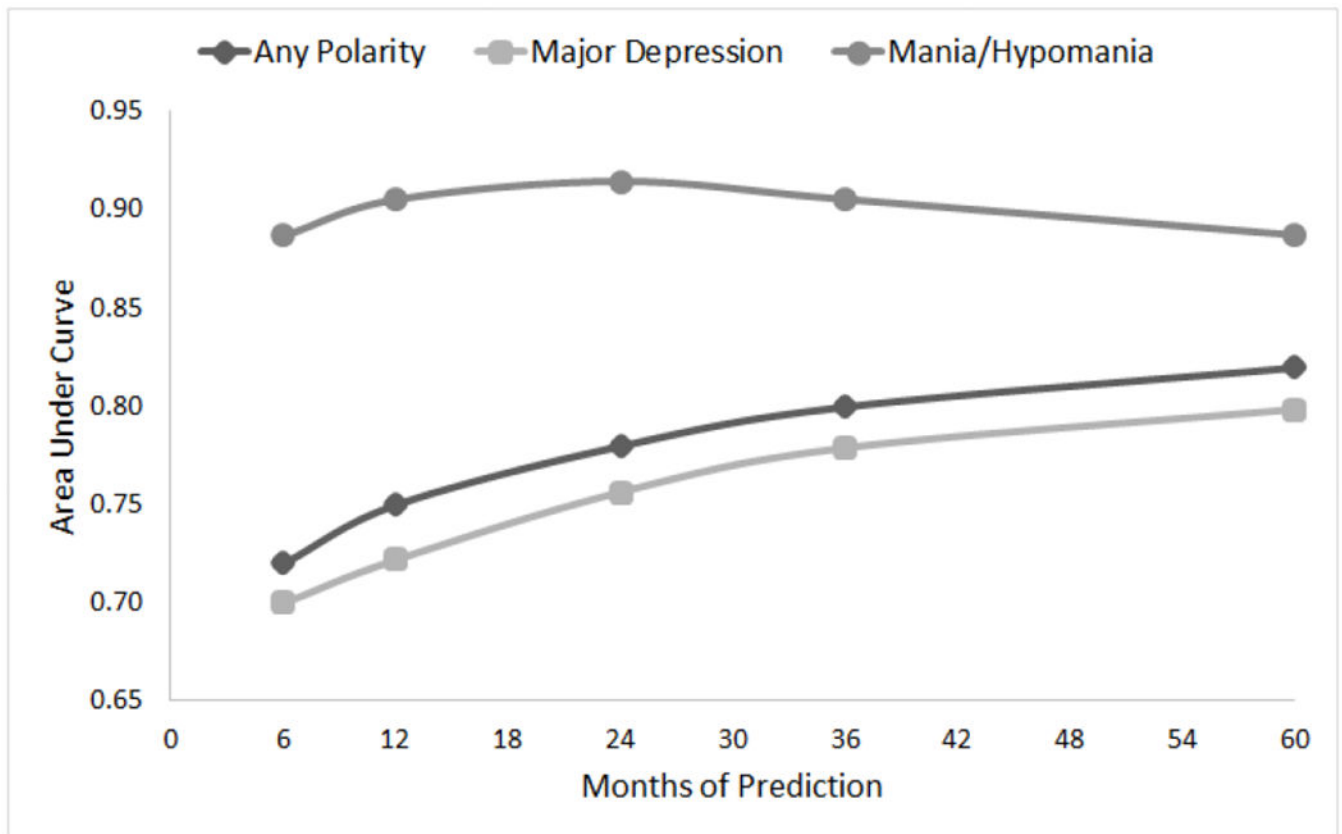


Figure 1. Area Under the Curve (AUC) by Length of Prediction Interval of Youths With Bipolar Disorder (Training: n=182, Testing: n=181)

Note: PSR = Psychiatric Status Ratings (PSR) from the Longitudinal Interval Follow-Up Evaluation (LIFE); PSR 1-2 = no or minimal symptoms, PSR 3-4 = subthreshold mood symptoms, PSR 5 = syndrome symptomatology; Recovery = 2 months with PSRs 1-2; Recurrence= PSRs 5 for at least 1 week for hypo/mania, and at least 2 weeks for a major depressive episode.

Table 1.

Demographic and Clinical Characteristics of Youths with Bipolar Disorder (Training: n=182, Testing: n=181)

Intake Demographic Variables	
Mean Age at Intake (SD)	12.6 (3.2)
Mean Age at the last follow-up (SD)	24.5 (4.6)
Mean Socioeconomic Status (SD)	3.4 (1.2)
Sex (% Female)	171 (47.1%)
Race (% Caucasian)	299 (82.4%)
Live with Both Biological Parents (%)	161 (44.4%)
Intake Clinical Variables	
Mean Age of Mood Disorder Onset (SD)	9.4 (3.9)
Bipolar Disorder (BD) Subtype	
BD-I	218 (60.1%)
BD-II	25 (6.9%)
BD-NOS	120 (33.1%)
Generalized Anxiety Disorder (%)	52 (14.3%)
ADHD (%)	208 (57.3%)
Disruptive Behavioral Disorders (%)	165 (45.5%)
Psychosis (%)	83 (22.9%)
Physical/Sexual Abuse (%)	72 (19.8%)
Polarity of the Index Episode (%)	
Depression	51 (14.0%)
Mania	69 (18.9%)
Hypomania	30 (8.2%)
Mixed	62 (17.0%)
Not Otherwise Specified	153 (41.9%)
Mean Total Number of New Recurrences (SD)	
Depression	2.3 (1.7)
Mania	1.3 (0.8)
Hypomania	1.5 (0.9)
Mixed	1.1 (0.3)
Lifetime (intake plus follow-up) Family History	
Depression (%)	322 (88.7%)
Mania/Hypomania (%)	211 (58.1%)
Anxiety (%)	269 (74.1%)
ADHD (%)	168 (46.3%)
Conduct Disorder (%)	126 (34.7%)
Psychosis (%)	64 (17.6%)
Substance Use Disorder (%)	254 (70.0%)
Suicidality (%)	188 (51.8%)

Note: Disruptive behavior disorders indicates oppositional defiant disorder and/or conduct disorder. Substance use disorder indicates all alcohol and substance abuse and dependence. ADHD = attention-deficit/hyperactivity disorder;

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Predictive Metrics for Range of Recurrence-Risk Score Cutoffs for the 5-year Risk Calculator for Youths With Bipolar Disorder (Training: n=182, Testing: n=181)

Predicted Risk Cutoff Decision Statistics (Any Recurrence)				
Risk Score Cutoff	Proportion of Sample in Risk Group	Sensitivity	Specificity	Positive Predictive Value
0.40	0.67	0.87	0.57	0.72
0.45	0.60	0.81	0.66	0.75
0.50	0.53	0.74	0.74	0.78
0.55	0.44	0.65	0.83	0.82
0.60	0.37	0.58	0.89	0.86

Optimal Sensitivity/Specificity = 0.74 at Cutoff = 0.50

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Individual Predictive Value of Each Factor Included in the Risk Calculator of Youths With Bipolar Disorder
(Training: n=182, Testing: n=181)

Predictor Domain	Factors	Across-Fold Average Relative Influence	Test AUC Decrement if Removed from Model		
			Any Recurrence	Depressive Recurrence	Manic Recurrence
Demographic/Family Factors	Age at assessment	16.23%	-0.025 ^a	-0.019	-0.010
	Age of mood disorder onset	7.31 %	-0.016 ^a	-0.021 ^a	-0.020
	Family History of Mania (1st vs. 2nd vs. Both)	2.09%	-0.003	-0.007	-0.004
Factors From the Previous Episode	Maximum PSR major depression score	10.36%	-0.007	-0.008	-0.007
	Number of weeks with threshold major depression	5.59%	-0.005	-0.007	-0.005
	Maximum PSR hypo/mania score	3.63%	0.000	-0.003	-0.002
	Number of weeks with threshold hypo/mania	1.56%	0.000	-0.005	-0.006
Recovery Factors	Current recovery length	15.84%	-0.009	-0.030 ^a	-0.005
	Prior recovery length	31.19%	-0.016	-0.014	-0.011
Past Episode History Factors	Number of Recurrences (None vs. 1 vs. 2+)	0.33%	0.000	-0.003	-0.003
	Episodes mostly include hypo/manic symptoms?	5.88%	0.000	-0.008	-0.012

Note: AUC = area under the receiver operating characteristic curve; PSR = psychiatric status ratings from the Longitudinal Interval Follow-Up Evaluation (LIFE); PSR 1-2 = no or minimal symptoms, PSR 3-4 = subthreshold mood symptoms, PSR 5 = syndrome symptomatology; Recovery = 2 months with PSRs 1-2; Recurrence = PSR 5 for at least 1 week for hypo/mania, and at least 2 weeks for a major depressive episode.

^aSignificantly different from zero after bootstrapping 95% CIs.

Table 4.

Additional Demographic and Clinical Factors Evaluated for the 5-year Risk Calculator for Youths With Bipolar Disorder (Training: n=182, Testing: n=181)

Additional Predictor Variables	Effect of Adding Individual Predictors		Effects of Adding Full Predictor Domain	
	AUC Increment	Relative Influence	AUC Increment	Relative Influence
Demographics				
Female	0.000	0.05		
Caucasian	0.001	0.52		
SES	0.000	0.92	0.002	2.44
Living with both biological parents	0.001	1.32		
Comorbid Diagnoses				
Any Comorbid Diagnosis	0.000	0.04		
GAD	0.000	0.00		
ADHD	0.000	0.06		
DBD	0.000	0.31	0.005	2.34
Psychosis	0.001	0.25		
SUD	0.004	1.78		
History of Physical/Sexual Abuse	0.000	0.18	0.000	0.18
Suicidality				
Lifetime Suicide Attempt	0.000	0.06		
Lifetime Threshold Suicidal Ideation	0.000	0.14	0.000	0.20
Expanded Family History				
Depression	0.000	0.64		
ADHD	0.001	1.11		
CD	0.002	1.38	0.011	7.56
Anxiety	0.000	1.02		
SUD	0.007	4.12		
Suicidality	0.003	1.44		
Most Severe Mood Symptoms Lifetime				
Major Depression	0.001	1.86	0.001	3.41
Hypo/mania	0.000	1.70		
Out-of-Episode Subthreshold Mood Symptoms (Presence vs. Absence)				
Subthreshold Major Depression	0.000	0.19	0.001	0.58
Subthreshold Hypomania	0.000	0.39		
Index Episode Polarity	0.002	2.00	0.002	2.00

Note: ADHD = attention-deficit/hyperactivity disorder; BD = bipolar disorder; DBD = disruptive behavior disorders (incorporating oppositional defiant disorder and/or conduct disorder); GAD = generalized anxiety disorder; PSR = Psychiatric Status Ratings from the Longitudinal Interval Follow-Up Evaluation (LIFE); PSR 1-2 = no or minimal symptoms, PSR 3-4 = subthreshold mood symptoms, PSR 5 = syndrome symptomatology; Recovery = 2 months with PSRs 1-2; Recurrence = PSRs 5 for at least 1 week for hypo/mania, and at least 2 weeks for a major depressive episode; SES = socioeconomic status; SUD = substance use disorder (incorporating all alcohol and substance abuse and dependence);

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript