

UC San Diego

UC San Diego Previously Published Works

Title

A risk calculator to predict suicide attempts among individuals with early-onset bipolar disorder

Permalink

<https://escholarship.org/uc/item/9t41k77x>

Journal

Bipolar Disorders, 24(7)

ISSN

1398-5647

Authors

Goldstein, Tina R
Merranko, John
Hafeman, Danella
[et al.](#)

Publication Date

2022-11-01

DOI

10.1111/bdi.13250

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



Published in final edited form as:

Bipolar Disord. 2022 November ; 24(7): 749–757. doi:10.1111/bdi.13250.

A Risk Calculator to Predict Suicide Attempts Among Individuals with Early-Onset Bipolar Disorder

Tina R. Goldstein, PhD¹, John Merranko, MA¹, Danella Hafeman, MD PhD¹, Mary Kay Gill, MSN¹, Fangzi Liao, MA¹, Craig Sewall, PhD¹, Heather Hower, MA², Lauren Weinstock, PhD², Shirley Yen, PhD^{2,3}, Benjamin Goldstein, MD PhD⁴, Martin Keller, MD², Michael Strober, PhD⁵, Neal Ryan, MD¹, Boris Birmaher, MD¹

¹Department of Psychiatry, Western Psychiatric Hospital, University of Pittsburgh School of Medicine, Pittsburgh PA

²Department of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University, Providence RI

³Massachusetts Mental Health Center and the Department of Psychiatry, Harvard Medical School at Beth Israel Deaconess Medical Center, Boston MA

⁴Centre for Addiction and Mental Health (CAMH), Toronto Ontario, Canada

⁵Department of Psychiatry, University of California, Los Angeles CA

Abstract

Objectives: To build a one-year risk calculator (RC) to predict individualized risk for suicide attempt in early-onset bipolar disorder.

Methods: 394 youth with bipolar disorder who completed ≥ 2 follow-up assessments (median follow-up length=13.1 years) in the longitudinal Course and Outcome of Bipolar Youth (COBY) study were included. Suicide attempt over follow-up was assessed via the A-LIFE Self-Injurious/Suicidal Behavior scale. Predictors from the literature on suicidal behavior in bipolar disorder that are readily assessed in clinical practice were selected and trichotomized as appropriate (presence past 6 months/lifetime history only/no lifetime history). The RC was trained via boosted multinomial classification trees; predictions were calibrated via Platt scaling. Half of the sample was used to train, and the other half to independently test the RC.

Results: There were 249 suicide attempts among 106 individuals. Ten predictors accounted for >90% of the cross-validated relative influence in the model (AUC=0.82; in order of relative influence): 1) age of mood disorder onset; 2) non-suicidal self-injurious behavior (trichotomized); 3) current age; 4) psychosis (trichotomized); 5) socioeconomic status; 6) most severe depressive symptoms in past 6 months (trichotomized none/subthreshold/threshold); 7) history of suicide attempt (trichotomized); 8) family history of suicidal behavior; 9) substance use disorder (trichotomized); 10) lifetime history of physical/sexual abuse. For all trichotomized variables, presence in the past 6 months reliably predicted higher risk than lifetime history.

Conclusions: This RC holds promise as a clinical and research tool for prospective identification of individualized high-risk periods for suicide attempt in early-onset bipolar disorder.

Keywords

suicide; bipolar disorder; risk calculator; child and adolescent

Introduction

Suicide is currently the second leading cause of death worldwide among youth aged 12–24.¹ The majority of youth who die by suicide meet criteria for at least one psychiatric disorder; of these, bipolar disorder confers greatest risk. Nearly one-third of youth with bipolar disorder attempt suicide at least once by mid-adolescence,² 60% attempt at least once in their lifetimes, and up to 20% will die by suicide, rendering the risk of death by suicide 60 times higher among individuals with bipolar disorder than that of the general population.³

Data from cross-sectional, prospective, and meta-analytic studies identify specific factors associated with elevated risk for suicidal behavior among individuals with bipolar disorder, one of the most potent of which is early age of bipolar illness onset.⁴ Others include mixed episodes, psychosis, non-suicidal self-injury, family history of suicide, and history of abuse.^{5–7} Identification of these group-level risk factors is critical, but yield information about risk for the group as a whole. As such, they may not reliably predict risk for any specific individual. Moreover, since these risk factors are both numerous and heterogenous, it is not surprising that a recent large-scale meta-analysis shows these factors predict risk for suicide attempt only slightly better than chance,⁸ as do clinical judgment,⁹ and clinical rating scales.¹⁰

Given that suicidal behavior results from a complex interaction of factors, suicide prevention efforts among high-risk groups like individuals with early-onset bipolar disorder could be enhanced by the capacity to predict an *individual's* risk for SA over a specified time period. A risk calculator (RC) is a clinical tool that uses an optimally identified set of risk factors to compute the probability of a specified event in an individual patient.¹¹ RCs are widely used by physicians to enhance clinical decision-making across several health conditions including stroke and cancer.^{12, 13} Recently, RCs have been developed to predict course and treatment response in psychiatric disorders including psychosis¹¹ and depression.¹⁴ Our group developed RCs to predict 5-year bipolar disorder onset among offspring of parents with bipolar disorder,¹⁵ risk for progression from bipolar-not otherwise specified (NOS) to bipolar-I/-II among youth,¹⁶ and risk for mood recurrence among young adults with bipolar disorder.¹⁷

Efforts to create RCs to predict individual-level risk for suicide attempt have been limited to date. Spittal et al.¹⁸ developed the 4-item Repeated Episodes of Self-Harm (RESH) scale to predict 6-month risk of selfharm repetition (with or without suicidal intent) among hospitalized individuals, with AUC=0.75. Fazel et al.¹⁹ describe the 17-item OxMIS (Oxford Mental Illness and Suicide Tool) to predict suicide death over 1 year among individuals with severe mental illness (AUC=0.71). These RCs demonstrate moderate predictive power (62–100% at varying thresholds and 75%, respectively), but low sensitivity (6–74% and 55%) among varied samples. Prediction may be improved by constructing tailored models within particular groups at higher risk, like individuals with early-onset bipolar disorder.

Furthermore, ultimate clinical benefit of RCs is contingent upon inclusion of predictors readily assessed across settings. To highlight, Melhem et al.²⁰ computed a risk score for suicide attempt over 12 years among offspring of parents with mood disorders using 6 readily assessed clinical factors. The model yielded high sensitivity (87%) and moderate improvement in specificity over the aforementioned tools (63%) but lacks their temporal specificity.

We describe development of a RC to predict an individual's one-year risk for SA for youth with bipolar disorder followed for a median of 13.1 years in the multi-site longitudinal Course and Outcome of Bipolar Youth (COBY) study. To enhance clinical utility, predictors of suicidal behavior in bipolar disorder readily assessed in clinical practice were identified from systematic reviews.

Methods

We present a brief summary of COBY study methods below. More details are in prior COBY publications.²¹

Participants

From October 2000 through July 2006, COBY recruited 446 youth ages 7–17 (mean intake age=12.6, $SD=3.3$, 53% male, 82% White) with a primary diagnosis of DSM-IV bipolar-I, -II, or COBY operationalized Not Otherwise Specified (-NOS; see Psychiatric Diagnosis). Youth with schizophrenia, developmental delay, and mood disorders secondary to substances, medications, or medical illness were excluded. To model the RC, we examined data from 394 participants who completed at least two follow-up assessments (median follow-up=13.1 years; range=1–18 years; Table 1); we assessed longitudinal data retrospectively at average follow-up periods of 36 weeks ($SD=18$ weeks).

Procedures

The Institutional Review Boards at the three participating sites approved all study procedures. Prior to study procedure administration, parents and participants provided informed consent. Assessments were conducted by trained master's level staff who maintained reliability on all study assessments.²¹

Psychiatric Diagnosis. Study assessors interviewed parents and children at study intake to determine current and past psychiatric diagnoses. We used the Schedule for Affective Disorders and Schizophrenia for School-Aged Children--Present and Lifetime Version (K-SADS-PL)²² to assess for non-mood psychiatric disorders, and the K-SADS-P mood sections,²³ plus additional items from the K-SADS Mania Rating Scale (MRS),²⁴ to assess mood symptoms. Consensus scores reflect data from youth and parents, and child psychiatrists/psychologists confirmed all diagnoses. KSADS-PL kappas ≥ 0.8 for psychiatric disorders, and K-MRS and KSADS-P depression section intraclass correlation coefficients (ICCs) ≥ 0.95 . Eligible participants met criteria for DSM-IV bipolar-I ($n=292$), -II ($n=56$), or -NOS ($n=65$), using research criteria operationalized for COBY.²¹ Age at mood disorder

onset was when the participant first met DSM-IV criteria for any mood episode (i.e., hypo/manic, mixed, major depressive, or bipolar-NOS).

Suicide Attempt. History of suicide attempt was defined as any self-injurious act that reached or exceeded an operationalized threshold (see below) of lethal intent and/or medical lethality, assessed via the K-SADS-P depression section suicidal acts items²³ (assessed during the current and most severe past episodes), and/or the K-SADS Summary Lifetime Diagnostic Checklist SA item. Past non-suicidal self-injury was assessed via the K-SADS-P non-suicidal self-damaging acts item (positive if ≥ 3 , infrequent). We utilized the A-LIFE Self-Injurious/Suicidal Behavior Scale²⁵ to document all self-injurious behaviors, *regardless of intent*, during follow-up. For each self-injurious behavior over the follow-up period, date, method, intent [using K-SADS Depression Scale item ratings:²³ 1 (none) to 6 (extreme, careful planning, and every expectation of death)] and medical threat [1 (no danger) to 7 (death)] are coded. In keeping with our prior work,²⁵ suicide attempt was defined as any self-injurious behavior that included definite intent (≥ 3 , definite but ambivalent) *and/or* mild lethality (≥ 3). All self-injurious behaviors for which intent and lethality were rated minimal or less (≤ 2) were considered non-suicidal self-injurious behavior.

Demographic Variables. We obtained demographic information via a parent-report form, and history of treatment and abuse via medical history questionnaire. Socioeconomic status (SES) was ascertained using the four-factor Hollingshead scale.²⁶ A demographic update form documented changes over follow-up.

Family Psychiatric History. Study assessors administered the Structured Clinical Interview for DSM-IV²⁷ with the participating parent at intake to gather parental psychiatric history; familial psychiatric history was collected via an enhanced version of the Family History Screen²⁸ and updated at each follow-up assessment. Family history for each condition (including lifetime suicide attempt) was considered positive if any first- or second-degree relative met criteria for ‘probable’ or ‘definite’ disorder.

Psychiatric Symptoms. We documented changes in psychiatric symptoms over follow-up using the Adolescent Longitudinal Interview Follow-Up Evaluation (A-LIFE).¹⁹ The A-LIFE Psychiatric Status Rating Scales (PSR) yield week-by-week symptom ratings using numeric values operationally linked to DSM-IV criteria, with demonstrated reliability and external validity.²¹ At each follow-up assessment, assessors retrospectively rated mood disorder severity using a 1–6 scale (1=no symptoms, 2–4=subthreshold DSM-IV symptoms, 5–6=full threshold DSM-IV criteria), and comorbid diagnoses on a 1–3 scale (1=no symptoms, 2=subthreshold, 3=full threshold). At intake, history of psychotic symptoms was “present” if either the hallucinations or delusions item on the K-SADS-PL screening interview was rated “threshold, definitely present,” or “moderate” or greater (≥ 4) on the K-MRS. Over follow-up, hallucinations and delusions were rated weekly on the A-LIFE PSR, and ratings of “definite” (3; 2=probable, 1=absent) were considered “present.” History of suicidal ideation at intake was considered positive if K-SADS-P depression summary scores reflected past or current suicidal ideation rated “mild” (3) or greater. Over follow-up, suicidal ideation was rated weekly on the A-LIFE PSR using the K-SADS suicidal

ideation 1–6 severity scale.²³ Mixed episodes over follow-up were operationalized as either a threshold (A-LIFE PSR 5/6) rating for both depression and hypo/mania during the same follow-up week, or a threshold depression or hypo/mania rating and a subthreshold rating on the opposite pole.²⁵ Lithium exposure was tracked weekly over follow-up using the A-LIFE Psychotropic Treatment Schedule. ICCs for A-LIFE syndromal/subsyndromal mood disorders were ≥ 0.75 .

Hospitalization. A medical history questionnaire documented history of psychiatric hospitalization at intake. Assessors document receipt and timing of psychosocial treatment over follow-up using the A-LIFE Psychosocial Treatment Schedule.

Physical/Sexual Abuse. History of physical and sexual abuse was assessed at intake and each follow-up via the K-SADS-PL Traumatic Events form.

Statistical Analysis

As in prior publications, we applied criteria developed by Fusar-Poli et al¹¹ to develop the RC. To ensure potential for widespread clinical utility of the resultant RC, we selected predictors from extant systematic reviews on risk factors for suicidal behavior in individuals with bipolar disorder that are readily assessed in clinical practice (Table 2). Strict reliance on our own work identifying predictors of suicidal behavior in COBY may have resulted in model overfitting. Risk factors were trichotomized into three timeframes: threshold presence in the past 6 months; lifetime history only; no lifetime history.

The RC was trained via boosted multinomial classification trees,²⁹ a useful model for these data because it implicitly incorporates interactions between predictors and has been shown to effectively predict mood episodes in COBY and elsewhere.^{17,30} Predictions were then calibrated via Platt scaling.³¹ To avoid overfitting, approximately half of the sample was used to train the RC; the other half was held out and used to independently test the RC. The training sample was further subdivided into five folds so that Platt scaling could be performed via cross-validation (~40 subjects and >500 observations per fold).³² We designed and implemented an algorithm to optimize balance between randomized sets and folds in suicide attempt rates, with randomization at the participant level (i.e., repeated measurements within-participant were confined to the same fold; see e-Methods 1). In addition to the boosted model, we trained and tested a parametric prediction model using ridge regression for comparison. Shrinkage parameters were determined using methods from Cule and De Iorio.³³

To test model validity, we assessed discrimination via area under the receiver operating characteristic curve (AUC) on the testing sample, evaluating the predicted one-year suicide attempt risk. Test sensitivity, specificity, and positive predictive value were assessed at a range of risk thresholds. Calibration was tested via Hosmer-Lemeshow test, and by plotting and comparing observed vs. predicted suicide attempt risk. Predictor importance was assessed via the Friedman relative influence statistic²⁹ (computed via cross-validation), and the decrement in AUC with removal of that variable from the model (significance of decrements tested via 95% bootstrap percentile interval). Lastly, an experimental reduced model was trained using only the ten predictors that comprise the top 90% of cross-validated

relative influence in the full model. Lastly, we examined whether predictors of suicide attempt risk similarly predicted the lethality of suicide attempts by reformulating the RC model to instead predict the medical lethality scores associated with each suicide attempt; cross-validated relative influence statistics were similarly calculated for each predictor variable.

Results

Rates of suicide attempt over follow-up.

Over a median of 13.1 years of follow-up, we observed 249 suicide attempts among 106 individuals. Twenty-seven percent of the sample (106/394) had at least one suicide attempt over follow-up, and 14% (54/394) had ≥ 2 suicide attempts. The estimated median time to first attempt over follow-up was 4 years (range=0.5–12 years). Among participants with a suicide attempt over follow-up, median rate was approximately one attempt per seven years.

One-Year Risk Calculator.

The participant randomization procedure yielded a training sample with 200 participants with 2751 observations, and a testing sample with 194 participants with 2776 observations. Except for a slightly higher rate of disruptive behavior disorder in the testing sample (63% vs. 51%, $p=0.01$), there were no significant between-subsample differences in length of follow-up, number of assessments, duration between assessments, number of suicide attempts, demographic or clinical factors. Overall, the participant randomization procedure yielded highly balanced subsamples.

The test AUC of the full model's one-year risk predictions was 0.82 (0.79, 0.85; Figure 1). Predicted and observed suicide attempt risk were consistent through the range of risk scores and did not significantly differ (Hosmer-Lemeshow $\chi^2=10.24$, $df=8$, $p=0.25$), indicating no evidence of miscalibration. Furthermore, average predicted one-year risk in the testing sample (mean=0.050, median=0.029) was similar to the observed suicide attempt rate in the testing sample (0.049). Table 3 shows test prediction metrics across a range of predicted risk thresholds; sensitivity and specificity are maximized at 0.73, with a predicted risk threshold of 0.036 (i.e., if predicted risk ≥ 0.036 , predict SA; if predicted risk < 0.036 , predict no suicide attempt).

Of the 21 predictors in the model (Table 4), 10 accounted for >90% of the cross-validated relative influence in the model (direction associated with increased risk noted in parentheses): 1) age of mood disorder onset (younger); 2) non-suicidal self-injury history (presence); 3) current age (older); 4) history of psychosis (present); 5) SES (lower); 6) most severe past 6 months depressive symptoms (i.e., PSR 1–6; higher); 7) history of suicide attempt (present); 8) family history of suicide attempt (present); 9) SUD history (present); and 10) lifetime history of physical/sexual abuse (present). Of note, when parametrically estimating the predictive influence of each level of a trichotomous variable, presence in the past 6 months reliably predicted higher suicide attempt risk than lifetime history.

When reformulating the model to predict medical lethality of suicide attempts, nine out of the ten most influential predictors in the RC outlined above remained among the top

ten most influential predictors of medical lethality (Supplemental Table 1). The notable exception was non-suicidal self-injury history, which, while highly predictive of risk for suicide attempt, was not highly predictive of medical lethality. This predictor was replaced by family history of bipolar disorder as one of the ten predictors comprising more than 90% of the cross-validated relative influence in the medical lethality model.

Sensitivity analyses indicated no significant AUC decrements after individually eliminating any predictors in the model, with the largest decrement = -0.03 (-0.04, 0.03) when removing age of mood disorder onset. Sensitivity analyses excluding youth with bipolar-II and -NOS (i.e., limited to bipolar-I only) yielded a minimal increase in AUC (0.83), with the same 10 predictors accounting for the top 90% of relative influence in the model. Similarly, sensitivity analyses excluding observations with concurrent lithium use, and separately including only those with concurrent lithium use, showed no significant change in AUC. Substituting dichotomous (i.e., presence in the past 6 months; lifetime history) for trichotomous risk factors in the model yielded nearly identical AUCs and influential variables, indicating minimal impact on model performance. Lastly, dividing substance use disorder history into two separate predictor variables for alcohol use disorder and other substance (i.e., drug) use disorder yielded a nearly identical AUC. Drug use disorder was a slightly more influential predictor than alcohol use disorder (4.54% relative influence vs. 1.89%). However, it is important to note that drug use disorder over follow-up was slightly more prevalent than alcohol use disorder in this sample (31% of participants vs. 26%), and within participants with substance use disorder, drug use disorder was much more persistent over follow-up than alcohol use disorder (27% of follow-up weeks on average vs. 16%).

The ridge regression trained and tested for comparison to the boosted model performed comparably weaker, with a test AUC=0.71 and maximized sensitivity and specificity at 0.65 with predicted risk threshold =0.0375. Predicted and observed suicide attempt risk were consistent through the range of risk scores and did not significantly differ (Hosmer-Lemeshow $\chi^2=6.58$, $df=8$, $p=0.58$), indicating no evidence of miscalibration.

A reduced model was trained using only the ten predictors accounting for the top 90% of cross-validated relative influence. Results indicated that model discrimination remained approximately unchanged in the reduced model at 0.82 (0.79, 0.84), and optimal sensitivity and specificity improved to 0.74.

Discussion

We describe the first individual-level 1-year RC for suicide attempt among the high-risk group of individuals with early-onset bipolar disorder. This RC offers promise for clinical use due to its reliance on just ten readily assessed demographic and clinical variables, including current age, age of mood disorder onset, SES, family history of suicide attempt, personal history of suicide attempt, non-suicidal self-injury and psychotic symptoms, and recent (past 6 months) threshold depressive symptoms. Additionally, sensitivity and specificity (AUC=0.82) of the RC meets or exceeds that of most RCs utilized in clinical practice for other medical (e.g., cardiovascular disease, AUC=0.76–0.79) and psychiatric conditions (e.g., new-onset psychosis, AUC=0.71).¹¹

Alternative approaches to prediction of suicidal behavior in high-risk groups include screening tools (e.g., the Comprehensive Suicide Risk Assessment for veterans³⁴) and self-reports (e.g., the Concise Health Risk Tracking Self-Report for adults with BP³⁵), but these have lower sensitivity and specificity for predicting near-term suicide attempt, and rely on patient-reported symptoms which may be subject to bias. Recent studies have attempted to enhance reliability and scalability of suicide risk prediction tools by applying machine learning algorithms to electronic medical record data,,e.g.,³⁶ yielding models with similar predictive power as this RC. While these models have the benefit of automated, large-scale risk detection in healthcare systems, they are limited to individuals for whom prior medical records are available. Thus, models dependent on electronic medical record data from within a single healthcare system may not be reliable or even implementable among individuals with mental health disorders, who are more likely to be uninsured and seek treatment from safety net providers.³⁷ Indeed, when de la Garza and colleagues³⁸ applied machine learning to predict suicide attempt in a representative sample of adults in the United States, socioeconomic disadvantage, lower education level and past year financial crisis were among the 10 most potent predictors, highlighting inherent behavioral healthcare disparities.

A clinical tool like the RC we describe relies on demographic and clinical information commonly gathered during standard psychiatric assessment, and as such, may hold promise to enhance clinical care. For example, results could trigger notifications alerting providers of individuals at elevated risk and changes in risk status. Elevated risk scores could further indicate the need for more detailed imminent suicide risk assessment using evidence-based tools, as well as creation, update, and/or review of the individual's safety plan.³⁹ Subsequent intervention may include consideration of appropriate frequency, type and level of care for the individual's predicted risk. Additionally, outreach efforts may be initiated for high-risk patients who do not attend scheduled visits, as implemented in the Veterans Affairs health system.⁴⁰ This approach is closely aligned with the aims of the American Foundation for Suicide Prevention's "Project 2025"⁴¹, a nationwide initiative to reduce the annual suicide rate 20% by 2025. This plan specifically highlights enhanced identification of at-risk individuals via evidence-based practices within systems of primary and behavioral healthcare as one of four priority areas with greatest potential for impact.

While the RC we describe yields highly accurate individual-level prediction, a RC score does not replace clinical judgment. Indeed, additional factors shown to predict elevated risk for suicidal behavior among youth with bipolar disorder—including several included herein—did not emerge as highly influential in the RC (e.g., sex at birth, mixed symptoms²⁵), it is possible that such factors exert their effects through the more highly predictive factors in the RC. These and other established risk factors (e.g., comorbid personality disorder⁵)⁵ should also be considered for select individuals with early-onset bipolar disorder. In some cases, the RC may under- or overestimate risk, and additionally does not yield detailed information on specific aspects of suicide risk (e.g., intent, lethality), rendering further clinical and suicide risk assessment critical. Of note, our analyses examining a secondary RC to predict suicide attempt lethality demonstrates remarkable predictor overlap with the RC to predict suicide attempt, with nine of the top ten most influential variables remaining the same. The exception is history of non-suicidal self-injury, which is highly predictive of

risk for suicide attempt, but not of medical lethality. Together, these findings indicate the RC holds promise as a data-driven tool to augment and guide suicide risk assessment and treatment planning.

Regarding research implications, the RC may help identify individuals during elevated periods of risk for whom novel suicide prevention approaches may be examined. Additionally, future studies may further enhance person-level prediction of near-term suicide risk by combining the RC with cognitive tasks and/or biological indices (e.g., neuroimaging, polygenic risk scores) implicated in suicide risk.¹⁰

Strengths of this study include the well-characterized sample of individuals with early-onset bipolar disorder followed longitudinally with validated measures. However, included risk factors were limited to those collected and therefore are not exhaustive (e.g., emotional abuse). Additional limitations include a primarily White sample, rendering examination among more diverse samples a critical future direction. COBY participants were recruited from clinical settings, possibly limiting generalizability of the results. Next steps include external validation in an independent sample of individuals with early-onset bipolar disorder, and examination among community samples and individuals with adult-onset bipolar disorder. Future work may also aim to identify models that reliably predict risk over even more proximal timeframes (e.g., one month).

Using just ten demographic and clinical factors commonly assessed in clinical practice, this individual-level RC for early-onset bipolar disorder reliably predicts one-year suicide attempt risk. This low-burden tool, if validated, offers promise to enhance suicide risk assessment and management in this high-risk population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

The authors wish to thank the youth and their family members who participated in the Course and Outcome of Bipolar Youth (COBY) study the COBY research team, and administrative support from Pamala Pyle. This research was supported by National Institute of Mental Health (NIMH) grants MH59929 (PI: Boris Birmaher, M.D.), MH59691 (PIs: Martin B. Keller, M.D., Shirley Yen, Ph.D.), and MH59977 (PI: Michael Strober, Ph.D.).

References

1. Curtin SC, Heron M. Death rates due to suicide and homicide among persons aged 10–24: United States, 2000–2017. Vol. 352. 2019. NCHS Data Brief. <https://www.cdc.gov/nchs/products/index.htm>.
2. Goldstein TR. Suicidality in pediatric bipolar disorder. *Child and Adolescent Psychiatric Clinics of North America*. 2009;18(2):339–352. [PubMed: 19264267]
3. Simon G, Hunkeler E, Firmina B, Lee J, Savarino J. Risk of suicide attempt and suicide death in patients treated for bipolar disorder. *Bipolar Disorders*. 2007;9:526–530. doi:10.1111/j.1399-5618.2007.00408.x [PubMed: 17680924]
4. Dome P, Rohmer Z, Gonda X. Suicide risk in bipolar disorder: A brief review. *Medicina*. 2019;55(8):403. doi:10.3390/medicina55080403 [PubMed: 31344941]

5. Schaffer A, Isometsa E, Tondo L, et al. International Society for Bipolar Disorders task force on suicide: meta-analyses and meta-regression of correlates of suicide attempts and suicide deaths in bipolar disorder. *Bipolar Disorders*. 2015;17(1):1–16. doi:10.1111/bdi.12271
6. da Silva Costa L, Alencar AP, Neto PJN, et al. Risk factors of suicide in bipolar disorder: a systematic review. *Journal of Affective Disorders*. 2015;170:237–254. doi:10.1016/j.jad.2014.09.003 [PubMed: 25261630]
7. Tidemalm D, Haglund A, Karanti A, Landen M, Runeson B. Attempted suicide in bipolar disorder: risk factors in a cohort of 6086 patients. *PLoS One*. 2014;9(4):e9409. doi:10.1371/journal.pone.0094097
8. Franklin JC, Ribeiro JD, Fox KR, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin*. 2017;143(2):187–232. doi:10.1037/bul0000084 [PubMed: 27841450]
9. Tran T, Luo W, Phung D, et al. Risk stratification using data from electronic medical records better predicts suicide risks than clinician assessments. *BMC Psychiatry*. 2014;14(76)doi:10.1186/1471-244X-14-76
10. Roos L, Sareen J, Bolton JM. Suicide risk assessment tools, predictive validity findings and utility today: time for a revamp? *Neuropsychiatry*. 2013;3(5):1–13. [PubMed: 28018489]
11. Fusar-Poli P, Rutigliano G, Stahl D, et al. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. *JAMA Psychiatry*. 2017;74(5):493–500. doi:10.1001/jamapsychiatry.2017.0284 [PubMed: 28355424]
12. Nobel L, Mayo NE, Hanley J, Nadeau L, Daskalopoulou SS. MyRisk_Stroke Calculator: A personalized stroke risk assessment tool for the general population. *Journal of Clinical Neurology*. 2014;10(1):1–9. doi:10.3988/jcn.2014.10.1.1 [PubMed: 24465256]
13. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being evaluated annually. *Journal of the National Cancer Institute*. 1989;81(24):1879–1886. doi:10.1093/jnci/81.24.1879 [PubMed: 2593165]
14. Perlis RH. A clinical risk stratification tool for predicting treatment resistance in major depressive disorder. *Biological Psychiatry*. 2013;74(1):7–14. doi:10.1016/j.biopsych.2012.12.007 [PubMed: 23380715]
15. Hafeman D, 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. doi:10.1001/jamapsychiatry.2017.1763 [PubMed: 28678992]
16. Birmaher B, Merranko J, 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. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2018;57(10):755–763. doi:10.1016/j.jaac.2018.05.023 [PubMed: 30274650]
17. Birmaher B, Merranko JA, Gill MK, et al. Predicting personalized risk of mood recurrences in youths and young adults with bipolar spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2020;59(10):1156–1164. doi:10.1016/j.jaac.2019.12.005 [PubMed: 31978620]
18. Spittal MJ, Pirkis J, Miller M, Carter G, Studdert DM. The Repeated Episodes of Self-Harm (RESH) score: A tool for predicting risk of future episodes of self-harm by hospital patients. *Journal of Affective Disorders*. 2014;161:36–42. doi:10.1016/j.jad.2014.02.032 [PubMed: 24751305]
19. Fazel S, Wolf A, Larsson H, Mallett S, Fanshawe TR. The prediction of suicide in severe mental illness: Development and validation of a clinical prediction rule (OxMIS). *Translational Psychiatry*. 2019;9(98)doi:10.1038/s41398-019-0428-3
20. Melhem NM, Porta G, Oquendo M, et al. Severity and variability of depression symptoms predicting suicide attempt in high-risk individuals. *JAMA Psychiatry*. 2019;76(6):603–612. [PubMed: 30810713]
21. Birmaher B, Axelson D, Strober M, et al. Clinical course of children and adolescents with bipolar spectrum disorders. *Arch Gen Psychiatry*. 2006;63(2):175–183. doi:10.1001/archpsyc.63.2.175 [PubMed: 16461861]

22. 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. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1997;36(7):980–988. doi:10.1097/00004583-199707000-00021 [PubMed: 9204677]
23. 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. doi:10.1001/archpsyc.1985.01790300064008 [PubMed: 4015311]
24. Axelson DA, Birmaher B, 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. *Journal of Child and Adolescent Psychopharmacology*. 2003;13(4):463–470. doi:10.1089/104454603322724850 [PubMed: 14977459]
25. Goldstein TR, Ha W, Axelson DA, et al. Predictors of prospectively observed suicide attempts among youth with bipolar disorder. *Archives of General Psychiatry*. 2012;69(11):1113–1122. doi:10.1001/archgenpsychiatry.2012.650 [PubMed: 22752079]
26. Hollingshead A Four-factor Index of Social Status. Yale University; 1975.
27. First MB, Gibbon M. The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). In: Hilsenroth MJ, Segal DL, eds. *Comprehensive Handbook of Psychological Assessment: Personality Assessment*. John Wiley & Sons, Inc.; 2004:134–143.
28. 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. NOT IN FILE. doi:10.1001/archpsyc.57.7.675 [PubMed: 10891038]
29. Friedman JH. Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*. 2001;29(5):1189–1232.
30. Van Voorhees BW, Paunesku D, Glollan J, Kuwabara S, Reinecke M, Basu A. Predicting future risk of depressive episode in adolescents: the Chicago Adolescent Depression Risk Assessment (CADRA). *Annals of Family Medicine*. 2008;6(6):503–511. doi:10.1370/afm.887 [PubMed: 19001302]
31. Niculescu-Mizil A, Caruana R. Obtaining calibrated probabilities from boosting. *Association for Uncertainty in Artificial Intelligence Press*; 2005:413–420.
32. Platt JA. Probabilities for SV Machines. In: Smola JA, Bartlett PL, Scholkopf B, Schuurmans D, eds. *Advances in large margin classifiers*. MIT Press; 2000:61–74.
33. Cule E, De Iorio M. Ridge regression in prediction problems: automatic choice of the ridge parameter. *Genet Epidemiol*. Nov 2013;37(7):704–14. doi:10.1002/gepi.21750 [PubMed: 23893343]
34. Doran N, De Peralta S, Depp C, et al. The Validity of a Brief Risk Assessment Tool for Predicting Suicidal Behavior in Veterans Utilizing VHA Mental Health Care. *Suicide and life-threatening behavior*. 2016;46(4):471–485. doi: 10.1111/sltb.12229 [PubMed: 26822821]
35. Reilly-Harrington NA, Shelton RC, Kamali M, et al. A tool to predict suicidal ideation and behavior in bipolar disorder: The Concise Health Risk Tracking Self-Report. *Journal of Affective Disorders*. 2016;192:212–218. doi:10.1016/j.jad.2015.12.036 [PubMed: 26748736]
36. Simon GE, Johnson E, Lawrence JM, et al. Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. *American Journal of Psychiatry*. 2018;175(10):951–961. doi:10.1176/appi.ajp.2018.17101167 [PubMed: 29792051]
37. Mechanic D, Olfson M. The relevance of Affordable Care Act for improving mental health care. *Annual Review of Clinical Psychology*. 2016;12:515–542.
38. de la Garza A, Blanco C, Olfson M, Wall M. Identification of suicide attempt risk factors in a national US survey using machine learning. *JAMA Psychiatry*. 2021;78(4):398–406. [PubMed: 33404590]
39. Stanley B, Brown G. Safety planning intervention: A brief intervention to mitigate suicide risk. *Cognitive and Behavioral Practice*. 2012;19(2):256–264. doi:10.1016/j.cbpra.2011.01.001

40. McCarthy JF, Bossarte RM, Katz IR, et al. Predictive modeling and concentration of the risk of suicide: Implications for preventive interventions in the US Department of Veterans Affairs. *American Journal of Public Health*. 2015;105(9):1935–1942. [PubMed: 26066914]
41. Prevention AFfS. A nationwide initiative to reduce the annual rate of suicide in the U.S. 20 percent by 2025. 2017. accessed 12/3/21. <https://project2025.afsp.org/>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

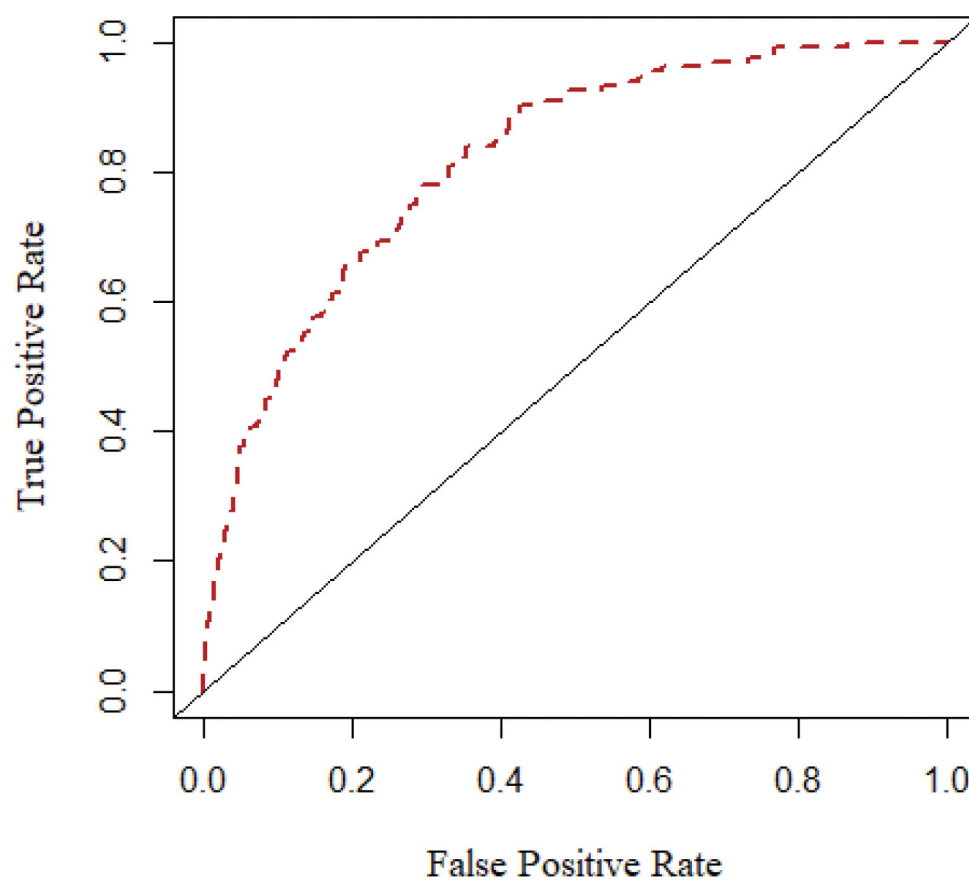


Figure 1.
Receiver Operating Characteristic Curve

Table 1.
Demographic and Clinical Characteristics of COBY Youth (N=394)

Follow-up Statistics	Median, Mean (SD)
Follow-up Length in Years	13.13, 12.43 (3.62)
Duration Between Assessments in Weeks	36, 40 (17.73)
Intake Variables	Mean (SD) or N (%)
Age	12.70 (3.28)
Socioeconomic Status (SES) [†]	3.41 (1.18)
Sex at birth (% female)	186 (47.21%)
Race (% White)	324 (82.23%)
Living with both biological parents (%)	167 (42.39%)
Age of mood disorder onset	9.32 (3.93)
BP subtype	
I	236 (59.90%)
II	28 (7.11%)
NOS [‡]	130 (32.99%)
Lifetime/Follow-Up Clinical Variables	N (%)
Suicide attempt	184 (46.70%)
Suicidal ideation	312 (79.19%)
Non-suicidal self-injurious behavior	208 (52.79%)
Psychiatric hospitalization	274 (69.54%)
Mixed episode	157 (39.85%)
Psychotic symptoms	146 (37.06%)
Sexual/physical abuse	186 (47.21%)
Comorbid disorders	
Anxiety disorder	293 (74.37%)
Substance use disorder	167 (42.39%)
Family history	
BP	230 (58.38%)
Depression	347 (88.07%)
Anxiety Disorder	290 (73.60%)
Substance use disorder	277 (70.30%)
Suicide attempt	171 (43.40%)

[†]Hollingshead-Redlich Criteria

[‡]COBY operationalized criteria

Abbreviations: BP = Bipolar Disorder; COBY = Course and Outcome of Bipolar Youth.

Table 2.

Predictors from Existing Literature ^{4,6,7} Included in the 1-Year Suicide Attempt Risk Calculator

Predictor
Age
Sex at birth
Socioeconomic status
Bipolar disorder subtype
Age of mood disorder onset
History of suicide attempt
History of non-suicidal self-injurious behavior
History of physical/sexual abuse
History of psychiatric hospitalization
Panic disorder
Any anxiety disorder
Substance use disorder
History of psychotic symptoms
History of mixed episodes
Depressive episode severity/burden
Family history of bipolar disorder
Family history of suicide attempt
Family history of substance use disorder

Abbreviation: COBY = Course and Outcome of Bipolar Youth

Table 3.
Predictive Metrics for Risk Score Cutoffs for 1-Year Suicide Attempt Risk Calculator

Predicted Risk Cutoff Decision Statistics				
Risk Score Cutoff	Proportion of Sample in Risk Group	Sensitivity	Specificity	Positive Predicted Value
0.030	0.46	0.91	0.56	0.10
0.035	0.31	0.77	0.71	0.12
0.040	0.23	0.67	0.79	0.14
0.045	0.18	0.58	0.84	0.16
0.050	0.14	0.53	0.88	0.18

Table 4.
Individual Predictive Value of Each Factor Included in the 1-Year Suicide Attempt Risk Calculator

Predictor Variable	Across-Fold Average Relative Influence	AUC Decrement if Removed from Model
Age of mood disorder onset	21.45	0.03
History of non-suicidal self-injury [†]	14.11	0.00
Current age	13.60	0.00
History of psychotic symptoms [†]	8.94	0.01
Socioeconomic status [‡]	8.28	0.00
Depressive episode severity/burden [§]	6.74	0.00
History of suicide attempt [†]	5.74	0.00
Family history of suicide attempt	5.35	0.00
History of substance use disorder [†]	4.05	0.00
Lifetime history of physical/sexual abuse	3.87	0.01
Family history of bipolar disorder	1.43	
Hospitalization in previous six months (y/n)	1.40	
Duration of threshold depression symptoms [¶]	1.20	
Threshold suicidal ideation history*	1.13	
Female (sex at birth)	1.01	
Panic disorder	0.54	
Family history of substance use	0.41	
Any anxiety disorder	0.39	
Social support over previous six months	0.35	
Bipolar disorder subtype (NOS vs. I/II)	0.02	
Mixed episode in the previous six months (y/n)	0.00	

Shaded cells comprise the top 90% of relative influence. Note that no individual predictor's removal from the model yielded a statistically significant AUC decrement after bootstrap.

[†] Present during previous six months vs. lifetime history only vs. no lifetime history

[‡] Hollingshead-Redlich (1975) criteria

[§] Most severe depression symptoms in the past 6 months (A-LIFE PSR 1–6)

[¶] Six months of continuous symptoms vs. intermittent symptoms vs. no symptoms

Abbreviation: AUC = Area Under Curve; COBY = Course and Outcome of Bipolar Youth; A-LIFE PSR = Adolescent-Longitudinal Interval Follow-Up Evaluation Psychiatric Status Ratings (Keller et al., 1987).