

UC San Diego

UC San Diego Previously Published Works

Title

Validation of a youth suicide risk calculator in an adult sample with bipolar disorder.

Permalink

<https://escholarship.org/uc/item/07m6q05c>

Authors

Iyengar, Satish

Hafeman, Danella

Hunt, Jeffrey

et al.

Publication Date

2024-02-01

DOI

10.1016/j.jad.2023.11.066

Peer reviewed



Published in final edited form as:

J Affect Disord. 2024 February 15; 347: 278–284. doi:10.1016/j.jad.2023.11.066.

Validation of a Youth Suicide Risk Calculator in an Adult Sample with Bipolar Disorder

Jess G. Fiedorowicz, M.D., Ph.D.^a, John A. Merranko, M.A.^b, Tina R. Goldstein, Ph.D.^b, Heather Hower, M.S.W.^{c,d,e}, Satish Iyengar, Ph.D.^f, Danella M. Hafeman, M.D., Ph.D.^b, Jeffrey I. Hunt, M.D.^{c,g}, Michael Strober, Ph.D.^h, Martin B. Keller, M.D.^{c,i}, Benjamin I. Goldstein, M.D., Ph.D.^j, Rasim S. Diler, M.D.^b, Sara Siddiqi, M.S.^a, Boris Birmaher, M.D.^b

^aDepartments of Psychiatry and Epidemiology, The University of Ottawa, 75 Laurier Ave. East, Ottawa, ON K1N 6N5, Canada

^bDepartment of Psychiatry, Western Psychiatric Hospital, University of Pittsburgh School of Medicine, 3811 O'Hara St., Pittsburgh, PA, 15213, USA

^cDepartment of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University, Box G-BH, Providence, RI, 02912, USA

^dDepartment of Health Services, Policy, and Practice, Brown University School of Public Health, 121 South Main Street, Providence, RI, 02903, USA

^eDepartment of Psychiatry, University of California San Diego, 4510 Executive Drive, Suite 315, San Diego, CA, 92121, USA

^fDepartment of Statistics, University of Pittsburgh, 230 S. Bouquet St., Pittsburgh, PA, 15213, USA

^gDepartment of Psychiatry, Bradley Hospital, 1011 Veterans Memorial Parkway, East Providence, RI, 02915, USA

^hDepartment of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

ⁱDepartment of Psychiatry, University of Miami, 1120 NW 14th St., Miami, FL, 33136, USA

^jDepartment of Psychiatry, CAMH, University of Toronto Faculty of Medicine, 2075 Bayview Ave., FG-53, Toronto, ON M4N-3M5, Canada

Abstract

Background: Bipolar disorder (BD) conveys the highest risk of suicide of all mental disorders. We sought to externally validate a risk calculator (RC) of suicide attempts developed in youth with BD from the Course and Outcome of Bipolar Youth (COBY) study in an adult sample.

Methods: A prospective cohort of adults with BD from the National Institute of Mental Health Collaborative Depression Study (CDS; N=427; mean (+/- SD) age at intake (36 +/- 13 years)) was secondarily analyzed to validate the COBY RC for one-year risk of suicide attempts/deaths.

Corresponding Author: Jess G. Fiedorowicz, M.D., Ph.D. **Address:** The Ottawa Hospital, General Campus, 501 Smyth Road, Box 400, Ottawa, ON K1H 8L6, Canada. jfiedorowicz@toh.ca.

Nine of the ten predictor variables from the COBY RC were available in the CDS and used: age, age of mood disorder onset, first and second (partial) degree family history of suicide, history of psychotic symptoms, substance use disorder, prior suicide attempt, socioeconomic status, and non-suicidal self-injury (prospectively, incompletely at baseline).

Results: Over a mean (SD) follow-up of 19(10) years, 29% of the CDS sample attempted suicide. The RC predicted suicide attempts/deaths over one-year follow-up with an area under the receiver operating characteristic curve (AUC) of 0.78 (95% CI 0.75–0.80). The RC performed slightly better in those with a younger age of mood disorder onset.

Limitations: Clinical samples may limit generalizability; the RC does not assess more acute suicide risk.

Conclusions: One-year risk of suicide attempts/deaths can be predicted with acceptable accuracy in youth and adults with BD, comparable to commonly used RCs to predict cardiovascular risk. This RC may help identify higher-risk individuals with BD for personalized treatment and research. <https://cobysuicideattemptsrc.shinyapps.io/Shiny>

Keywords

Bipolar disorder; Personalized medicine; Prospective cohort study; Risk calculator; Suicide attempts

1. Introduction

Risk calculators (RCs) are clinical tools that use an optimally identified set of predictors to compute the probability of some specified event for a specific patient. An abundance of RCs have been validated and implemented for varied aspects of medical practice (e.g., cancer, cardiovascular disease (CVD), surgery) (1). While many of these RCs have found their way into other areas of medical practice, psychiatry has been arguably slow to adopt quantitative measures of individual risk into clinical practice (2). Of the models that have been developed in psychiatry, only about 20% underwent external validation in an independent sample, and those assessing clinical utility have focused only on predicting risk of developing a psychotic disorder (3).

There are few psychiatric RCs focused on bipolar disorder (BD) (<6% compared to >24% for psychotic disorders, >9% for post-traumatic stress disorder, >34% for depression) (3), which is striking given the well-described heterogeneity in the course of illness of BD (4, 5). Our group has developed four RCs to assess the individual prognosis of individuals with BD: 1) a 5-year RC for estimating risk of BD onset among offspring of parents with BD (6); 2) a 5-year RC for progression from BD not-otherwise specified (BD-NOS) to bipolar I disorder (BD-I) or bipolar II disorder (BD-2) in youth (7); 3) a 6-month to 5-year RC for mood recurrence in youth (8) with BD that was externally validated in an adult sample (9); and 4) a 1-year RC for suicide attempts in youth with BD (10).

Suicide represents an outcome for which risk estimation is of great interest, especially in BD, which is associated with approximately 20–30 times the general population risk of suicide, the highest of mental disorders (11). Suicide risk assessment is a routine part of

psychiatric practice, and this largely relates to the assessment of acute suicide risk (12). Validated tools for the assessment of acute risk remain elusive. However, tools for more intermediate risk assessment do provide some value to both clinical practice and research and their development has been more promising. Goldstein et al. developed a RC using data from the Course of Outcome of Bipolar Youth (COBY) study to estimate one-year risk of suicide attempts among individuals with early-onset BD (10). This RC included 10 predictors selected from the existing literature and was built from a training sample. In the holdout (test) sample, the calculator was able to predict one year risk of suicide attempts with an area under the receiver operating characteristic curve (AUC) of 0.82 with a sensitivity of 37% at 95% specificity. This level of accuracy may be adequate for cost-effective suicide risk prevention strategies based on statistical models in U.S. primary care of suicide risk calculators (13) and the accuracy is on par with other risk prediction models routinely used in medicine (14).

Ultimately, external validation is required before suicide RCs can be brought into clinical practice. The baseline and longitudinal data collection in COBY was comparable to that of the National Institute of Mental Health Collaborative Depression Study (CDS). Our group has previously used CDS data to externally validate a RC for recurrence of mood episodes that was derived in COBY (9). Prior analyses of the CDS cohort have identified several risk factors for suicide attempts or death including depressive symptoms, male sex, hopelessness, prior suicide attempts, psychotic symptoms, and substance abuse (15–18). For this analysis, we sought to externally validate the Goldstein et al. one-year suicide RC that was developed from the COBY data (10) using CDS data, in a substantially older sample of individuals diagnosed with BD between 1978 and 1981.

2. Method

2.1 Participants

The prospective cohort studies, COBY and CDS, enrolled youth and adults respectively and have been described in more detail elsewhere (19, 20). From 2000 to 2006, COBY initially enrolled 413 youth with BD who met Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) (21) criteria for BD-I, BD-II, or operationalized COBY criteria for BD-NOS (see (19), for BD-NOS criteria). For the development of the suicide RC from COBY, a sample of 394 participants who completed at least two follow-up visits was utilized, with half the sample used to train the RC, and the other half held out to test the RC independently (10). The CDS enrolled a clinical sample of participants with mood disorders, including those with BD-I and BD-II, between 1978 and 1981 (22, 23). For the current external validation analyses, a sample of 427 BD-I and BD-II CDS participants were utilized.

Institutional Review Board / Research Ethics Board approval was obtained from each participating institution in COBY and CDS. In COBY, informed consent/assent was obtained from participants and their parents. In CDS, written informed consent was obtained from participants.

2.2 Measures

CDS participants had a baseline assessment with the Schedule for Affective Disorders and Schizophrenia (SADS) (24) and the Personal History of Depressive Disorders (23). Depressive episode severity was captured on a 6-point ordinal scale via the Longitudinal Interval Follow-up Evaluation (LIFE) using Psychiatric Status Ratings (PSRs) (25). Age of mood disorder onset was obtained from the SADS using Research Diagnostic Criteria (RDC). The presence of psychotic symptoms and substance use disorder was further assessed at both baseline and follow-up. Socioeconomic status was assessed using the five-point Hollingshead-Redlich scale (26). Non-suicidal self-injury (NSSI) was not assessed at baseline for those with a history of suicide attempts. History of suicide attempt at baseline was collected in the SADS and thereafter assessed prospectively in follow-up. The SADS included an item for lifetime suicide attempt severity. If this item met the lowest two severity ratings for both intent and actual medical threat, this was classified as NSSI at baseline. For intent, this would correspond to, “obviously no intent, purely manipulative” or “not sure or only minimal intent” and, for threat, this would correspond to “no danger (e.g., no effects, held pills in hand)”, or “minimal (e.g., scratch on wrist)”. NSSI was captured over follow-up using the same measures, consistent with the methods of Goldstein et al. (10). A family history of suicide attempts was obtained from the family history study which has been detailed elsewhere (27, 28). While COBY assessed family history in first- and second-degree relatives, in the CDS, family history was obtained only in first-degree relatives and half-siblings. Other second-degree relative data were not available.

2.3 Statistical Methods

We trained the previously published COBY youth suicide attempts RC with the 10 predictor variables: 1) age of mood disorder onset; 2) NSSI; 3) current age; 4) psychosis; 5) socioeconomic status (SES); 6) most severe depressive symptoms in the past 6 months; 7) history of suicide attempt; 8) family history of suicidal behavior; 9) substance use disorder (SUD); and 10) lifetime history of physical/sexual abuse. We utilized 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 (8, 30). We then calibrated one-year risk predictions via Platt scaling (31). To avoid overfitting, we split the COBY sample into approximately equal training and testing sets. The training set was then used to train the RC, and the testing set was used to independently test the RC. The training set was further subdivided into five folds so that Platt scaling could be performed via cross-validation (~40 participants 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).

We then externally validated the COBY RC using longitudinal data from CDS, estimating AUC for one-year predictions of suicide attempt risk. Of note, as the CDS dataset did not include history of physical/sexual abuse, this variable could not be entered into the RC when estimating risk of suicide attempts in CDS, so 9 out of 10 predictor variables were available. Fortunately, the model used to train the RC is designed to accommodate

missing predictor values, meaning there was no need to retrain the RC. Sensitivity analyses assessed performance: 1) stratified by age of onset; and 2) restricting the outcome to suicide death. An exploratory, post hoc analysis assessed the RC score over time in those who died by suicide compared to a subsample of at least five age matched comparison participants per case. Model reduction was explored using feature reduction algorithms (greedy backward selection and hill climb) to simplify the RC, training models with 7 or fewer predictors on COBY and then evaluating AUC and sensitivity at 95% specificity in the CDS. The 95% specificity threshold was selected for comparison to the work of Ross et al (13). Models were trained on the COBY sample and selected based on performance in the CDS sample to avoid overfitting. The risk calculator can be found at <https://cobysuicideattempts.shinyapps.io/Shiny/>.

3. Results

3.1 CDS External Validation Sample Characteristics

The CDS external validation sample included 427 participants with a prospective diagnosis of BD-I or BD-II (mean age at intake 36.3 ± 13.2 ; 58% female; 66% BD-I). Participants were followed for a median of 22 years. Over follow-up, we observed a total of 346 suicide attempts and 18 suicide deaths in the CDS sample. Relevant sociodemographic and clinical characteristics are shown in Table 1.

3.2 Comparison Between CDS External Validation and COBY Sample Characteristics

As compared to the COBY sample, the CDS external validation sample was much older on average (age at intake 36.3 vs. 12.7 years, $p < 0.001$) and with substantially later mean ages of mood disorder onset (23.9 vs. 9.3 years, $p < 0.001$). CDS also featured a greater representation of female participants and less frequent substance use disorder, NSSI, and family history of suicide attempts as compared to COBY (Table 1, p -values < 0.04). Further, the CDS sample was followed nearly twice as long as the COBY sample (median follow-up durations = 22 vs. 13 years). Despite these differences, CDS and COBY participants had comparable rates of suicide attempts both over follow-up (28.6% vs. 26.9%, respectively) and lifetime (52.2% vs. 46.7%, respectively).

3.3 CDS External Validation Sample Risk Calculator

The RC showed good prediction of suicide attempts over the subsequent year, with an AUC of 0.78 (95% CI 0.75 – 0.80) in the CDS external validation sample, compared to 0.82 from COBY. The Receiver Operator Characteristic (ROC) curves from the CDS external validation and the validation of the COBY holdout sample are illustrated in Figure 1. In the CDS external validation sample, at a specificity of 95% (as noted above selected based on the work of Ross et al. (13)), the sensitivity was 29%, the positive predictive value was 14%, and the negative predictive value was 98%. For the validation of the COBY holdout sample, again with 95% specificity, sensitivity was 37%, positive predictive value was 28%, and negative predictive value was 97%. As shown in Figure 2, calibration of CDS predictions was good among observations from the lower half of the estimated risk distribution; however, predictions among observations from the upper half of the estimated risk distribution were underestimated, as demonstrated by observed risk exceeding predicted

risk. To test the robustness of the RC, we individually removed each predictor from the model, as shown in Table 2, and re-estimated AUCs for predictions in the CDS external and COBY holdout samples. AUC decrements ranged from -0.01 to -0.10 , with the largest decrement resulting from the removal of major depression symptom severity over the previous six months. AUCs were more robust to removal of predictor variables when validating on the COBY holdout sample compared to the CDS external validation sample except for removing age of mood disorder onset, which had a slightly smaller impact on the AUC in the CDS external validation sample compared to the COBY holdout sample (AUC decrements: -0.01 vs. -0.03).

3.4 CDS External Validation Sample Risk Calculator Sensitivity Analyses

Given that the COBY sample included only participants with early age of mood disorder onset, we conducted a sensitivity analysis to see how the RC performed on the CDS external validation sample separately among those with an early-onset (<21 years of age, $n=195$) vs. late-onset (≥ 21 years of age, $n=232$) mood disorder. RC performance was slightly better in the early-onset subsample (which more closely matched the COBY sample) than in the late-onset subsample (AUCs: 0.79 vs. 0.75). When specifically assessing the RC's ability to discriminate between suicide deaths and nonevents, the AUC is unchanged (0.78). Exploratory analysis of those who died by suicide showed on average their RC scores steadily increased over time approaching death. The RCs of age matched controls slightly decreased over time, as shown in Supplemental Figure 1.

3.5 Model Reduction

The correlations between predictor variables in both CDS and COBY samples are reported in Supplemental Figure 2. A reduced model requiring only five predictors (current age, depression severity, history of NSSI, history of suicide attempts, and substance use disorder) was identified using a greedy backward selection algorithm which demonstrated AUCs of 0.77 in COBY and 0.78 in CDS as shown in Supplemental Table 1. At 95% specificity, this model demonstrated 29% sensitivity in COBY and 26% sensitivity in CDS.

4. Discussion

This analysis externally validated the one-year RC for suicide attempt that was developed in a high-risk sample with early-onset BD. Nine of the ten predictor variables from the COBY youth RC were available in the CDS and used: 1) age of mood disorder onset; 2) NSSI; 3) current age; 4) psychosis; 5) SES; 6) most severe depressive symptoms in the past 6 months; 7) history of suicide attempt; 8) family history of suicidal behavior; and 9) SUD. Two items were only partially captured. NSSI was available at baseline only for those without prior suicide attempts and family history of second-degree relatives was limited to half-siblings. This external validation is also notable for being conducted in an adult sample from two decades earlier. While the RC performed slightly better for those with early-onset BD, the performance of the sample overall approximates that used in practice for other medical conditions. For example, in CVD, external validation of the American College of Cardiology / American Heart Association Pooled Cohort Risk Equation for atherosclerotic CVD (33) demonstrated an AUC of 0.74 in predicting 5-year risk (34) and

external validation of the 2008 Framingham 5-year CVD RC (35) identified AUCs of 0.76–0.77 (36). With a sensitivity of 29% at 95% specificity (again selected for comparison to work of Ross et al.), our model reaches the threshold estimated to provide good healthcare economic value in clinical practice (13). Other points on the curve could be selected, for example, with higher sensitivity and lower specificity. The model can be generated with missing variables and appears fairly robust to any single missing variable, with exception perhaps of severity of mood symptoms in the CDS, absence of which had less impact on the COBY sample.

In the CDS sample, in which suicide was more frequently observed than COBY (18 vs. 2 deaths), the RC was similarly able to predict suicide death as an outcome. A model reduced to five key variables also demonstrates sufficient accuracy. In the CDS sample, the RC was also reasonably calibrated apart from some underestimation of risk for the highest risk groups. Those who died by suicide also had steadily increasing risk scores, which was particularly prominent in the decade before death, and in stark contrast to the declining risk scores of a cohort of age-matched comparison participants. This suggests that the trajectory of the risk score from this RC may also be valuable in identifying those at highest risk of suicide mortality.

This RC is distinct from other validated suicide RCs based on its focus on BD and the number of predictors. The Oxford Mental Illness and Suicide (OxMIS) tool uses 17 predictors, compared to 10 predictors in the current RC. In an external validation sample with schizophrenia spectrum and BD, the OxMIS tool predicted suicide deaths at one year with an AUC of 0.70 (95% CI 0.69–0.71) (37), compared to the AUC of 0.78 in the current study for suicide attempts and deaths. The Durham Risk Score consists of a 23-item checklist. Although not yet externally validated, this risk score was derived from a development sample that combined data from three U.S. prospective cohort studies and then validated in a holdout sample from those cohorts. The Durham Risk Score performed with considerable accuracy, predicting three-year risk of suicide attempts with AUCs approximating 0.9 (38). Sequeira et al. (39) developed a RC using electronic medical record data from an academic mental health hospital. Initial validation revealed an accuracy of 0.73 for death by suicide (sensitivity 22.9%) and 0.67 for suicide attempts (sensitivity 20.5%) at 4 years. Sensitivity for suicide attempt remained low (22.3%) on follow-up evaluation with two years of additional data. With missing data in the electronic medical record, scores could only be calculated for 38.5% of patients for suicide death and 45.5% for suicide attempts (39). The electronic medical record-based calculator of Walsh et al. fared better with prospective validation and recalibration in more general medical patients with an AUC of 0.80 across the academic medical center, although with a much lower AUC of 0.54 in behavioral health for 30-day risk of suicide attempt (40). To date, accuracy has been unsatisfactory for electronic medical record-based measures in mental health settings.

A one-year suicide risk score offers several potential clinical, research, and economic benefits. One-year risk is easy to communicate to patients and families and therefore to interpret. It also represents a common duration for the follow-up after clinical trials (e.g., open-label extension) (41) or for longer-term trials, providing important insights into the enduring benefits or harms of a treatment. Specifically for suicide attempts, between 15–

25% of people repeat self-harm within one year of an index episode, and present to the same hospital (42). A one-year risk score could be useful for identifying highest risk individuals for high-intensity, resource intensive interventions (e.g., day hospital program, dialectical behavioral therapy, and longer term outpatient outreach to patients and families). In a recent economic evaluation by Ross et al., a specificity of 95% and sensitivity of 25% from a primary care-based suicide risk prediction model could reduce suicide death rates by 1.6 per 100,000 person years if it is used to target cognitive behavioral therapy, based on an estimated one-year intervention benefit (13). If utilized in a similar fashion, the current RC could allow health systems to direct care in a manner that could result in a meaningful reduction in suicide.

Strengths of the current analysis include the external validation of the RC with a distinctly different sample with BD. The RC had also been previously validated in a holdout COBY sample distinct from the derivation COBY sample. This suggests that the RC likely generalizes to other samples with BD, although as noted, calibration may be necessary if being used to estimate absolute, instead of simply relative risk, as was necessary for the highest risk portions of this CDS external validation sample. The CDS sample was also well-phenotyped and followed up for ascertainment of outcome. The risk calculator is easy to use and publicly available at <https://cobysuicideattemptsrc.shinyapps.io/Shiny>.

4.1 Limitations

Some limitations of the current study must be noted. The CDS external validation was racially homogeneous and external validation in racially and ethnically diverse samples is needed. Data for NSSI was incomplete at baseline for those with a history of suicide attempts and there is some risk for misclassification with our threshold for NSSI, although it was consistent to that of Goldstein et al. (10) Second degree family history data was also limited to half-siblings and therefore incompletely ascertained. Information was not available related to abuse histories. It is nonetheless noteworthy that the RC performed well despite missing or incomplete data for these three variables. This suggests the RC may be robust to missing variables, as was shown in the original COBY validation study. This was also seen in the CDS external validation study for all but the severity of the most severe depressive symptoms in the past three months.

5. Conclusions

In summary, despite application to a much older sample with a later onset of mood disorder from a different period and limited ability to assess some variables, we externally validated a one-year suicide RC using nine of ten variables. The included variables are easily accessible clinically. The performance of the RC was within a range of accuracy for RCs used clinically in other areas of medicine even following variable reduction to just five variables. This demonstrates that intermediate-term (one-year) suicide risk in BD can be assessed within an acceptable range. While more acute suicide risk prediction remains a future goal, this one-year RC may be useful for targeted clinical interventions or for research seeking to enrich samples for higher suicide risk. Future research could also consider integration of these clinical variables with biological variables to improve prediction (43–45).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors wish to thank the COBY/PROBY and CDS studies' participants and their families, and the associated research teams. The authors thank Mary Kay Gill for her assistance. This research was supported by the National Institute of Mental Health (NIMH), Course and Outcome of Bipolar Youth (COBY) study, grants RO1 MH059929 (PI: Birmaher), RO1 MH59691 (PIs: Keller/ Yen), and RO1 MH59977 (PI: Strober), and Predicting Adult Outcomes in Bipolar Youth (PROBY) study, grants RO1 MH112544 (PI: Birmaher), and RO1 MH5270580 (PI: Yen).

REFERENCES:

1. Ankerst DP, Hoefler J, Bock S, Goodman PJ, Vickers A, Hernandez J, et al. Prostate Cancer Prevention Trial risk calculator 2.0 for the prediction of low- vs high-grade prostate cancer. *Urology*. 2014 Jun;83(6):1362–7. [PubMed: 24862395]
2. Fusar-Poli P, Hijazi Z, Stahl D, Steyerberg EW. The Science of Prognosis in Psychiatry: A Review. *JAMA Psychiatry*. 2018 Dec 01;75(12):1289–97. [PubMed: 30347013]
3. Meehan AJ, Lewis SJ, Fazel S, Fusar-Poli P, Steyerberg EW, Stahl D, et al. Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges. *Mol Psychiatry*. 2022 Jun;27(6):2700–8. [PubMed: 35365801]
4. Judd LL, Akiskal HS, Schettler PJ, Endicott J, Maser J, Solomon DA, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch Gen Psychiatry*. 2002 Jun;59(6):530–7. [PubMed: 12044195]
5. Judd LL, Akiskal HS, Schettler PJ, Coryell W, Endicott J, Maser JD, et al. A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar II disorder. *Arch Gen Psychiatry*. 2003 Mar;60(3):261–9. [PubMed: 12622659]
6. Hafeman DM, Merranko J, Goldstein TR, Axelson D, Goldstein BI, Monk K, et al. Assessment of a Person-Level Risk Calculator to Predict New-Onset Bipolar Spectrum Disorder in Youth at Familial Risk. *JAMA Psychiatry*. 2017 Aug 01;74(8):841–7. [PubMed: 28678992]
7. Birmaher B, Merranko JA, Goldstein TR, Gill MK, Goldstein BI, Hower H, 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 Oct;57(10):755–63.e4. [PubMed: 30274650]
8. Birmaher B, Merranko JA, Gill MK, Hafeman D, Goldstein T, Goldstein B, et al. Predicting Personalized Risk of Mood Recurrences in Youths and Young Adults With Bipolar Spectrum Disorder. *J Am Acad Child Adolesc Psychiatry*. 2020 Oct;59(10):1156–64. [PubMed: 31978620]
9. Fiedorowicz JG, Merranko JA, Iyengar S, Hower H, Gill MK, Yen S, et al. Validation of the youth mood recurrences risk calculator in an adult sample with bipolar disorder. *J Affect Disord*. 2021 Dec 01;295:1482–8. [PubMed: 34563392]
10. Goldstein TR, Merranko J, Hafeman D, Gill MK, Liao F, Sewall C, et al. A risk calculator to predict suicide attempts among individuals with early-onset bipolar disorder. *Bipolar Disord*. 2022 Nov;24(7):749–57. [PubMed: 36002150]
11. Miller JN, Black DW. Bipolar Disorder and Suicide: a Review. *Curr Psychiatry Rep*. 2020 Jan 18;22(2):6. [PubMed: 31955273]
12. Weber AN, Michail M, Thompson A, Fiedorowicz JG. Psychiatric Emergencies: Assessing and Managing Suicidal Ideation. *Med Clin North Am*. 2017 May;101(3):553–71. [PubMed: 28372713]
13. Ross EL, Zuromski KL, Reis BY, Nock MK, Kessler RC, Smoller JW. Accuracy Requirements for Cost-effective Suicide Risk Prediction Among Primary Care Patients in the US. *JAMA Psychiatry*. 2021 Jun 01;78(6):642–50. [PubMed: 33729432]
14. Badawy MAEM, Naing L, Johar S, Ong S, Rahman HA, Tengah DSNA, et al. Evaluation of cardiovascular diseases risk calculators for CVDs prevention and management: scoping review. *BMC Public Health*. 2022 Sep 14;22(1):1742. [PubMed: 36104666]

15. Fiedorowicz JG, Leon AC, Keller MB, Solomon DA, Rice JP, Coryell WH. Do risk factors for suicidal behavior differ by affective disorder polarity? *Psychol Med.* 2009 May;39(5):763–71. [PubMed: 18667100]
16. Coryell W, Kriener A, Butcher B, Nurnberger J, McMahon F, Berrettini W, et al. Risk factors for suicide in bipolar I disorder in two prospectively studied cohorts. *J Affect Disord.* 2016 Jan 15;190:1–5. [PubMed: 26476155]
17. Persons JE, Coryell WH, Solomon DA, Keller MB, Endicott J, Fiedorowicz JG. Mixed state and suicide: Is the effect of mixed state on suicidal behavior more than the sum of its parts? *Bipolar Disord.* 2018 02;20(1):35–41. [PubMed: 28833953]
18. Fawcett J, Scheftner W, Clark D, Hedeker D, Gibbons R, Coryell W. Clinical predictors of suicide in patients with major affective disorders: a controlled prospective study. *Am J Psychiatry.* 1987 Jan;144(1):35–40. [PubMed: 3799837]
19. Birmaher B, Axelson D, Strober M, Gill MK, Valeri S, Chiappetta L, et al. Clinical course of children and adolescents with bipolar spectrum disorders. *Arch Gen Psychiatry.* 2006 Feb;63(2):175–83. [PubMed: 16461861]
20. Rice J, Andreasen NC, Coryell W, Endicott J, Fawcett J, Hirschfeld RM, et al. NIMH Collaborative Program on the Psychobiology of Depression: clinical. *Genet Epidemiol.* 1989;6(1):179–82. [PubMed: 2731705]
21. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed., Text Revision ed. Washington, DC2000.
22. Katz MM, Klerman GL. Introduction: overview of the clinical studies program. *Am J Psychiatry.* 1979 Jan;136(1):49–51. [PubMed: 31801]
23. Katz MM, Secunda SK, Hirschfeld RM, Koslow SH. NIMH clinical research branch collaborative program on the psychobiology of depression. *Arch Gen Psychiatry.* 1979 Jul;36(7):765–71. [PubMed: 36864]
24. Endicott J, Spitzer RL. A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry.* 1978 Jul;35(7):837–44. [PubMed: 678037]
25. Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott P, et al. The Longitudinal Interval Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry.* 1987 Jun;44(6):540–8. [PubMed: 3579500]
26. Miller D The Hollingshead-Redlich scale. *Handbook of research design and social measurement.* White Plains, NY: Longman; 1983.
27. Fiedorowicz JG, Endicott J, Leon AC, Solomon DA, Keller MB, Coryell WH. Subthreshold hypomanic symptoms in progression from unipolar major depression to bipolar disorder. *Am J Psychiatry.* 2011 Jan;168(1):40–8. Epub 2010/11/17. [PubMed: 21078709]
28. Andreasen NC, Rice J, Endicott J, Coryell W, Grove WM, Reich T. Familial rates of affective disorder. A report from the National Institute of Mental Health Collaborative Study. *Arch Gen Psychiatry.* 1987 May;44(5):461–9. Epub 1987/05/01. [PubMed: 3579497]
29. Friedman J Greedy function approximation: a gradient boosting machine. *The Annals of Statistics.* 2001;29(5):1189–232.
30. 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):503–11. [PubMed: 19001302]
31. Niculescu-Mizil A, Caruana R. Obtaining calibrated probabilities from boosting. *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence.* 2005:413–20.
32. Platt J Probabilities for SV Machines. In: Smola J, Bartlett P, Scholkopf B, Schuurmans D, editors. *Advances in large margin classifiers:* MIT Press; 2000. p. 61–74.
33. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol.* 2014 Jul 1;63(25 Pt B):2889–934. Epub 2013/11/19. [PubMed: 24239923]

34. Rana JS, Tabada GH, Solomon MD, Lo JC, Jaffe MG, Sung SH, et al. Accuracy of the Atherosclerotic Cardiovascular Risk Equation in a Large Contemporary, Multiethnic Population. *J Am Coll Cardiol*. 2016 May 10;67(18):2118–30. Epub 2016/05/07. [PubMed: 27151343]
35. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation*. 2008 Feb 12;117(6):743–53. Epub 20080122. eng. [PubMed: 18212285]
36. Wallisch C, Heinze G, Rinner C, Mundigler G, Winkelmayr WC, Dunkler D. External validation of two Framingham cardiovascular risk equations and the Pooled Cohort equations: A nationwide registry analysis. *Int J Cardiol*. 2019 May 15;283:165–70. Epub 20181105. eng.
37. Sariaslan A, Fanshawe T, Pitkänen J, Cipriani A, Martikainen P, Fazel S. Predicting suicide risk in 137,112 people with severe mental illness in Finland: external validation of the Oxford Mental Illness and Suicide tool (OxMIS). *Transl Psychiatry*. 2023 Apr 18;13(1):126. Epub 20230418. eng. [PubMed: 37072392]
38. Kimbrel NA, Beckham JC, Calhoun PS, DeBeer BB, Keane TM, Lee DJ, et al. Development and validation of the Durham Risk Score for estimating suicide attempt risk: A prospective cohort analysis. *PLoS Med*. 2021 Aug;18(8):e1003713. Epub 20210805. eng. [PubMed: 34351894]
39. Sequeira L, McNair D, Wiljer D, Strudwick G, Deluca V, Kailasam K, et al. Implementing Predictive Models Within an Electronic Health Record System: Lessons from an External Validation of a Suicide Risk Model. *Stud Health Technol Inform*. 2022 Jun 06;290:562–6. [PubMed: 35673079]
40. Walsh CG, Johnson KB, Ripperger M, Sperry S, Harris J, Clark N, et al. Prospective Validation of an Electronic Health Record-Based, Real-Time Suicide Risk Model. *JAMA Netw Open*. 2021 Mar 01;4(3):e211428. Epub 20210301. eng. [PubMed: 33710291]
41. Megan B, Pickering RM, Weatherall M. Design, objectives, execution and reporting of published open-label extension studies. *J Eval Clin Pract*. 2012 Apr;18(2):209–15. Epub 20101010. eng. [PubMed: 21040252]
42. Zahl DL, Hawton K. Repetition of deliberate self-harm and subsequent suicide risk: long-term follow-up study of 11,583 patients. *Br J Psychiatry*. 2004 Jul;185:70–5. [PubMed: 15231558]
43. Roy B, Ochi S, Dwivedi Y. Potential of Circulating miRNAs as Molecular Markers in Mood Disorders and Associated Suicidal Behavior. *Int J Mol Sci*. 2023 Feb 28;24(5). Epub 20230228. eng.
44. Jackson NA, Jabbi MM. Integrating biobehavioral information to predict mood disorder suicide risk. *Brain Behav Immun Health*. 2022 Oct;24:100495. Epub 20220810. eng. [PubMed: 35990401]
45. Monson ET, Shabalin AA, Docherty AR, DiBlasi E, Bakian AV, Li QS, et al. Assessment of suicide attempt and death in bipolar affective disorder: a combined clinical and genetic approach. *Transl Psychiatry*. 2021 Jul 07;11(1):379. Epub 20210707. eng. [PubMed: 34234108]

ROC Curves: CDS vs. COBY

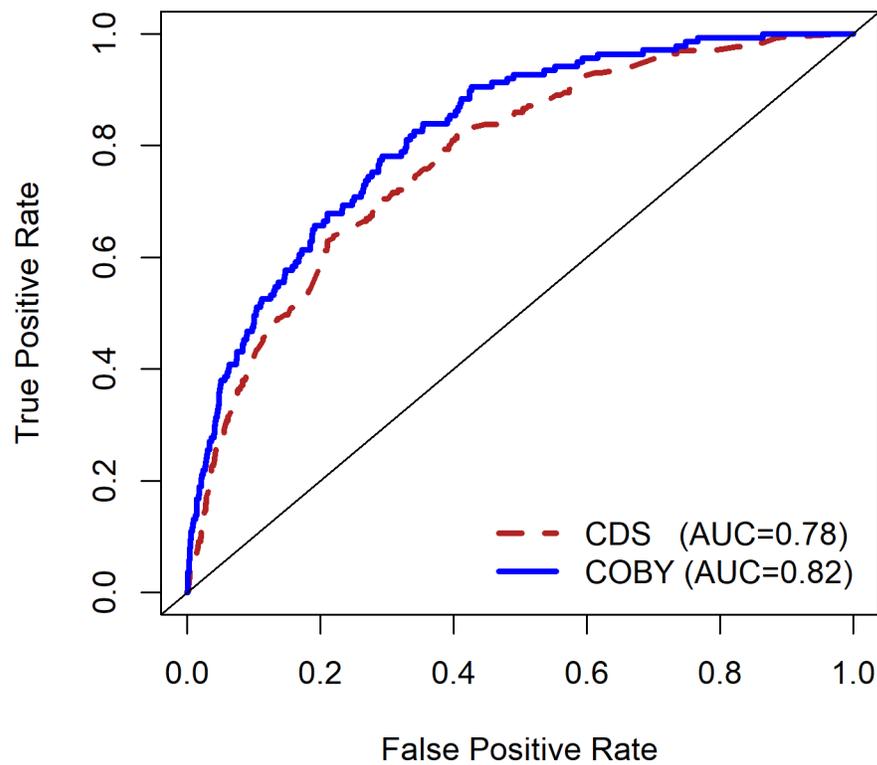


Figure 1. Receiver operating characteristic (ROC) curves for one-year predication of suicidal behavior.

The ROC curves compare the classification ability of the risk calculator in the CDS external validation sample (AUC=0.78) compared to the prior COBY holdout sample (AUC=0.82). AUC = Area Under the Curve, CDS = Collaborative Depression Study, COBY = Course and Outcome of Bipolar Youth study.

Calibration Curves: CDS vs. COBY

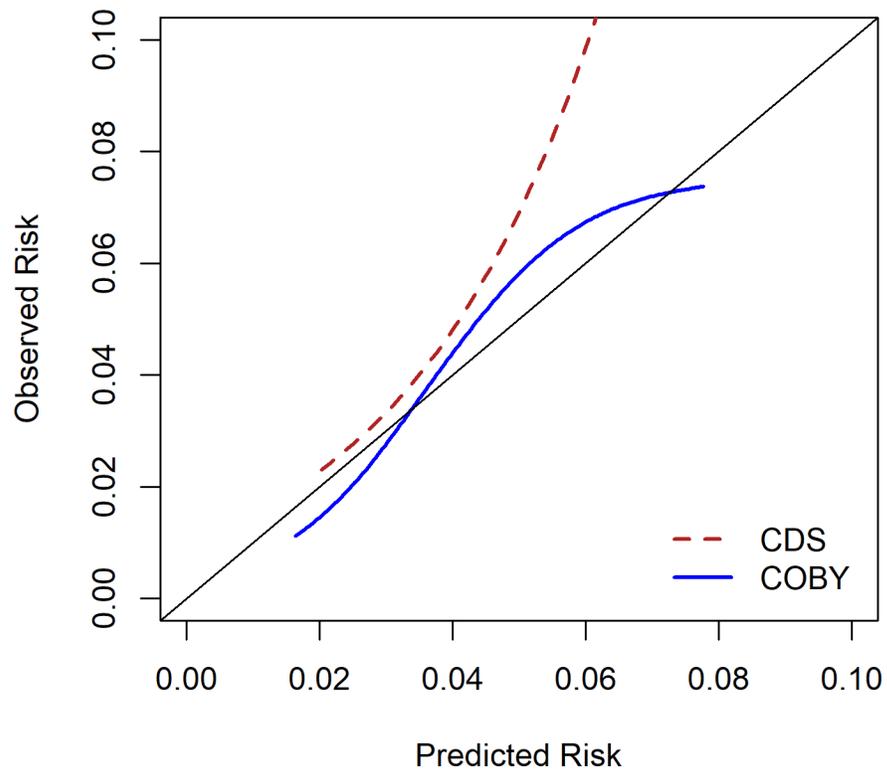


Figure 2. Calibration plot for five-year test predictions.

Predicted versus observed one-year suicide risk are displayed for the CDS external validation and the COBY holdout sample. The risk calculator underestimated risk in the higher risk portion of the sample in the CDS external validation cohort. CDS = Collaborative Depression Study, COBY = Course and Outcome of Bipolar Youth study.

Table 1.

Sample Comparisons of Demographic and Clinical Characteristics

Variable	CDS (N=427)	COBY (N=394)	Test Stat	p-value
Intake Age, mean (SD)	36.3 (13.2)	12.7 (3.3)	t=35.78	<0.001
Female, %	58.1%	47.2%	$\chi^2=9.72$	0.002
Socioeconomic Status, mean (SD)	3.4 (1.1)	3.4 (1.2)	t=0.10	0.92
Intake Diagnosis				
% Bipolar I Disorder	65.6%	59.9%	$\chi^2=213.69$	<0.001
% Bipolar II Disorder	34.4%	7.1%		
% BD Not-Otherwise-Specified	0.0%	33.0%		
Age of Mood Onset, mean (SD)	23.9 (10.2)	9.3 (3.9)	t=27.60	<0.001
Psychosis, %	42.6%	37.1%	$\chi^2=2.65$	0.10
Substance Use Disorder, %	35.4%	42.4%	$\chi^2=4.26$	0.04
Non-suicidal Self-injury over Follow-up, %	16.6%	30.0%	$\chi^2=20.52$	<0.001
Suicide Attempt over Follow-up, %	28.6%	26.9%	$\chi^2=0.28$	0.59
Lifetime History of Suicide Attempt, %	52.2%	46.7%	$\chi^2=2.50$	0.11
Family History of Suicide Attempt, %	2.6%	43.4%	$\chi^2=197.49$	<0.001

This table compares the sample from the CDS external validation to the original COBY sample from which the RC was derived. BD = bipolar disorder, CDS = Collaborative Depression Study, COBY = Course and Outcome of Bipolar Youth Study, SD = standard deviation

Table 2.

Impact on Area Under the Curve (AUC) from Removing Predictor Variables

Predictor Variable	Change to AUC if Removed from Model	
	COBY Holdout Sample	CDS External Validation Sample
Depressive episode severity/burden ^c	0.00	-0.10
History of suicide attempt ^a	0.00	-0.05
History of psychotic symptoms ^a	0.01	-0.02
Current age	0.00	-0.02
History of substance abuse ^a	0.00	-0.02
Age of mood disorder onset	-0.03	-0.01
Family history of suicide attempt	0.00	-0.01
Socioeconomic status ^b	0.00	-0.01
History of non-suicidal self-injury ^a	0.00	-0.01
Lifetime history of physical/sexual abuse	-0.01	---

^aPresent during previous six months vs. lifetime history only vs. no lifetime history

^bHollingshead-Redlich (1975) criteria

^cMost severe depression symptoms in the past 6 months (A-LIFE PSR 1-6)

A-LIFE = Adolescent Longitudinal Interval Follow-Up Evaluation, CDS = Collaborative Depression Study, COBY = Course and Outcome of Bipolar Youth study, PSR = Psychiatric Status Ratings