Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multidomain Childhood Bipolar Risk Clusters and Adolescent Bipolar Spectrum Emergence in the ABCD Study

Sam Sievertsen

sievertsen@ohsu.edu

Oregon Health & Science University https://orcid.org/0009-0002-2657-4902

Sara Shao

Scott Jones

Matthew Sullivan

Oregon Health & Science University https://orcid.org/0000-0002-2487-3936

Alexander Dufford

Oregon Health & Science University

Dani Del Rubin

Anna Wilson

Amy Holley

Bonnie Nagel

Rebekah Huber

Oregon Health & Science University https://orcid.org/0000-0002-6718-9683

Article

Keywords:

Posted Date: June 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9862629/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors have declared there is **NO** conflict of interest to disclose

Abstract

Bipolar disorder (BD) often emerges during adolescence with an average seven-year diagnostic delay associated with substantial burden and mortality. Determining whether BD risk factors form interpretable subgroups in childhood may inform earlier BD risk stratification and diagnosis. This study sought to identify how multidomain childhood BD risk factors cluster in a population-based cohort, and test whether clusters were prospectively associated with adolescent BD emergence. Leveraging baseline, 2-, 4-, and 6-year follow-up data from 11,868 youth in the Adolescent Brain Cognitive Development study, baseline risk features spanning clinical/prodromal, familial, neurocognitive, and environmental domains were clustered using k-prototypes. Primary outcomes were BD-I, BD-II, BD not otherwise specified, and any bipolar spectrum disorder (BSD), as derived from the KSADS-COMP. Pooled logistic discrete-time hazard models estimated BD onsets, and generalized estimating equations estimated BD prevalence over time. Complete baseline risk data from 9,193 participants yielded a stable 2-cluster solution, comprised of higher (n = 1,391) and lower (n = 7,802) risk clusters. Higher-risk membership was associated with greater BSD onset odds, (OR = 2.74, 95% CI [2.21,3.41], p< .001). By year-6, cumulative BSD risk was 19.7% in the higher-risk cluster versus 7.9% in the lower-risk cluster (RR = 2.50, 95% CI [2.06,3.03], p< .001). Higher-risk membership was also associated with greater BSD occurrence across follow-ups. In summary, multidomain childhood BD risk variables showed a stable 2-cluster profile that prospectively stratified adolescent BSD. Nearly 1 in 5 youths in the higher-risk cluster developed a BSD by mid-adolescence. This empirically derived framework provides a foundation for future BD risk stratification and external validation.

Introduction

Bipolar disorder (BD) often emerges during adolescence[1], a developmental period marked by rapid biological, social, and environmental change that can complicate diagnostic clarity[2]. The average delay from first onset of BD symptoms to accurate diagnosis is approximately 7-years[2]. Diagnostic ambiguity likely contributes to this delay, as early BD symptoms overlap with other psychiatric presentations and may initially manifest in non-specific ways. During this delay in receiving accurate diagnosis, affected youth may experience delayed or incorrect treatment, contributing to greater symptom severity, functional impairment, reduced quality of life[3,4], interference with educational achievement, meaningful social relationships, and increased familial stress and dysfunction[5]. Therefore, earlier detection of heightened risk for the development of BD, and BD itself, is imperative.

Relatedly, individual early-life risk factors associated with later BD manifestation have been identified, including depression and anxiety symptoms[6], attention problems[7], aggression[8], elevated mania symptoms[6,9], and greater negative and positive impulsivity[10]. However, many of these features are also shared across other forms of youth psychopathology, which limits the discriminative value of any single risk factor in isolation. Thus, while individual indicators may signal elevated BD vulnerability, they may be more informative when considered as part of broader multidomain risk profiles.

Beyond clinical and behavioral factors, family history of depression[11] or mania[12] may represent underlying genetic vulnerability, which may be compounded by environmental correlates of BD, including limited opportunity and resources[13], neighborhood safety problems[14], and bullying[15]. Sleep disturbance is also recognized as a core precursor and feature of BD, with early-onset BD cases especially prone to persistent sleep problems[16]. Neurocognitive factors such as working memory deficits, inhibitory control difficulties, and slower processing speed have also been documented in youth with BD or at familial risk for BD, and may represent early markers of vulnerability[17,18]. Lastly, exposure to adverse childhood experiences (ACEs) is associated with earlier BD onset and more severe illness course[19,20]. Considered together, extant findings suggest that childhood risk for BD may be better characterized by a constellation of risk factors across multiple domains than by any single indicator alone.

While many of these risk factors have been thoroughly investigated in longitudinal studies such as the Pittsburgh Bipolar Offspring Study[21] and the Course and Outcome of Bipolar Youth study[22], relatively few have evaluated how multiple domains of BD risk cluster during childhood prior to the emergence of the disorder. Prior seminal work has examined risk indicators individually or within specific domains, such as familial liability or early symptoms[23], and has focused primarily on predicting later BD emergence. However, BD has also been conceptualized as arising through a multidomain developmental process involving the interaction of clinical, familial, neurocognitive, and environmental factors[24]. As such, understanding the profile of childhood risk; that is, how factors in these risk domains cluster or co-occur before illness onset, represents a critical and underexamined area of investigation in characterizing BD's developmental trajectory[24].

This multidomain perspective is also consistent with Duffy's (2014) developmental clinical staging model of BD[25], which posits that early internalizing/externalizing symptoms and sleep dysregulation emerge on a background of familial liability, and are accompanied by neurocognitive vulnerabilities prior to the onset of syndromal illness. Identifying empirically derived configurations of these risk indicators in childhood may therefore provide insight into earlier developmental pathways of BD. Accordingly, guided by previously identified BD risk factors, research priorities outlined by the International Society for Bipolar Disorders pediatric BD task force[24], and Duffy's (2014) developmental clinical staging model[25], we a priori defined four staging-aligned domains of childhood BD risk in the current study: 1) clinical/prodromal; 2) familial; 3) neurocognitive; and 4) environmental/contextual.

These domains are readily assessed by the Adolescent Brain Cognitive Development (ABCD) study, the largest study of child health and brain development, following 11,868 children to understand how biological, environmental, and behavioral factors impact brain development and mental health through adolescence. By following a diverse population-based cohort from late childhood into mid-adolescence, the ABCD study provides a unique opportunity to evaluate whether staging-aligned, multidomain risk clusters at ages 9–10 (i.e., baseline) precede both the timing and burden of adolescent BD. This population-based design is particularly well suited to characterizing how multidomain BD risk profiles are distributed and cluster in the broader population, complementing prior work from high-risk cohorts[21,22].

More specifically, ABCD's breadth of baseline assessment and biennial outcome tracking enables data-driven identification of childhood risk clusters, with sufficient power to then examine associations between risk clusters and BD spectrum in adolescence. Therefore, in the current study, we leverage ABCD data to derive empirically defined risk clusters across these four risk domains, and test their association with the temporal onset and occurrence of adolescent BD. This approach aims to enhance our understanding of BD's early developmental trajectory, inform risk stratification and more targeted BD screening protocols, and highlight critical periods for intervention.

Methods

Dataset and Participants

The current study utilizes data from ABCD release 6.1 (https://doi.org/10.82525/ab4q-qr87). The ABCD Study initially recruited 11,868 participants beginning in 2016 from schools across 21 sites in the United States, sampled to ensure demographic representativeness of the urbanicity, socioeconomic status, race, ethnicity, and sex-at-birth of the country's population[26,27]. Participants included in the final analytic sample were selected based on completeness of baseline risk variables (n = 9,193) and the self-administered Kiddie Schedule for Affective Disorders and Schizophrenia Computerized Version (KSADS-COMP) for DSM-5 diagnoses and covariate data required for each longitudinal analysis. For models analyzing the onset or occurrence of BD, this included baseline BD diagnostic status and at least one eligible follow-up assessment at year-2, year-4, or year-6 (Table 1).

Table 1
Characteristics of Participants Included in the Primary Onset and Occurrence Analyses

Variable	Year-2	Year-4	Year-6
A. First Onset (DTH) Analytic Sample			
N Participants	8,396	7,139	3,883
Mid-interval age in years, Mean ± SD	11.02 ± 0.63	13.13 ± 0.68	15.20 ± 0.70
Sex at birth – Female, n (%)	3,958 (47.1%)	3,333 (46.7%)	1,804 (46.5%)
Race/ethnicity – Asian, n (%)	144 (1.7%)	118 (1.7%)	53 (1.4%)
Race/ethnicity – Black, n (%)	1,035 (12.3%)	839 (11.8%)	396 (10.2%)
Race/ethnicity – Hispanic, n (%)	1,619 (19.3%)	1,386 (19.4%)	701 (18.1%)
Race/ethnicity – Other[1], n (%)	836 (10.0%)	706 (9.9%)	375 (9.7%)
Race/ethnicity – White, n (%)	4,757 (56.7%)	4,086 (57.2%)	2,357 (60.7%)
Interval onset – BD-I, n (%)	48 (0.6%)	93 (1.3%)	57 (1.4%)
Interval onset – BD-II, n (%)	21 (0.3%)	36 (0.5%)	21 (0.5%)
Interval onset – BD-NOS, n (%)	54 (0.6%)	90 (1.3%)	60 (1.5%)
Interval onset – Any BSD, n (%)	95 (1.1%)	212 (3.0%)	124 (3.2%)
B. Occurrence (GEE) Analytic Sample			
N Participants	8,552	7,339	4,072
Age in years, Mean ± SD	12.07 ± 0.67	14.18 ± 0.71	16.08 ± 0.66
Sex at birth – Female, n (%)	4,019 (47.0%)	3,415 (46.5%)	1,900 (46.7%)
Race/ethnicity – Asian, n (%)	147 (1.7%)	123 (1.7%)	61 (1.5%)
Race/ethnicity – Black, n (%)	1,076 (12.6%)	885 (12.1%)	421 (10.3%)
Race/ethnicity – Hispanic, n (%)	1,646 (19.2%)	1,420 (19.3%)	734 (18.0%)
Race/ethnicity – Other[1], n (%)	852 (10.0%)	728 (9.9%)	396 (9.7%)

Note: The table depicts the characteristics of the two analytic cohorts across ABCD timepoints at follow-up years 2, 4, and 6. Panel **A** summarizes the first-onset analytic sample, and panel **B** summarizes the occurrence analytic sample. Participants were assigned to risk clusters using baseline risk variables (see results for details). [1] "Other" race category includes participants who identified as Native Hawaiian, Pacific Islander, Alaskan Native, American Indian, or multiracial. Abbreviations: BD-I = bipolar I disorder; BD-II = bipolar II disorder; BD-NOS = other specified bipolar disorder; BSD = bipolar spectrum disorder; DTH = discrete time hazard model; GEE = generalized estimating equation; SD = standard deviation.

Variable	Year-2	Year-4	Year-6
A. First Onset (DTH) Analytic Sample			
Race/ethnicity – White, n (%)	4,825 (56.4%)	4,179 (56.9%)	2,459 (60.4%)
Occurrence – BD-I, n (%)	56 (0.7%)	98 (1.3%)	60 (1.5%)
Occurrence – BD-II, n (%)	23 (0.3%)	38 (0.5%)	22 (0.5%)
Occurrence – BD-NOS, n (%)	64 (0.7%)	96 (1.3%)	65 (1.6%)
Occurrence – Any BSD, n (%)	109 (1.3%)	230 (3.1%)	143 (3.5%)

Note: The table depicts the characteristics of the two analytic cohorts across ABCD timepoints at follow-up years 2, 4, and 6. Panel **A** summarizes the first-onset analytic sample, and panel **B** summarizes the occurrence analytic sample. Participants were assigned to risk clusters using baseline risk variables (see results for details). [1] "Other" race category includes participants who identified as Native Hawaiian, Pacific Islander, Alaskan Native, American Indian, or multiracial. Abbreviations: BD-I = bipolar I disorder; BD-II = bipolar II disorder; BD-NOS = other specified bipolar disorder; BSD = bipolar spectrum disorder; DTH = discrete time hazard model; GEE = generalized estimating equation; SD = standard deviation.

BD Risk Measures and Variables

In the current study, 17 BD risk variables were selected from baseline assessment measures spanning clinical/prodromal, familial, neurocognitive, and environmental domains (Fig. 1). Clinical and behavioral measures reported by parents included the following four items from the Child Behavior Checklist (CBCL): DSM-5-oriented T-scored scales for depression and anxiety, and T-scored scales for attention problems and aggressive behavior[28]. Additional parent-reported behavioral measures included the sum scores from the Sleep Disturbance Scale for Children[29] and from the General Behavior Inventory (GBI) Mania scale[30]. CBCL and Sleep Disturbance Scale measures pertained to the previous 6 months, and GBI score pertained to lifetime behavior. Youth-reported measures included lifetime negative and positive urgency items from the abbreviated Urgency, Premeditation, Perseverance, Sensation seeking, Positive urgency Impulsive Behavior Scale[31].

Familial risk was indexed from parent-reported Family History Module items adapted for the ABCD study, which denote a lifetime history of depression or mania in first-degree relatives[32]. Environmental risk was assessed through the geospatially-derived current Child Opportunity Index 2.0[33], parent-reported perceived neighborhood safety from the PhenX Toolkit[34], and lifetime endorsement of bullying exposure from the KSADS-COMP, reported separately by both parents and youth[35]. Neurocognitive functioning was assessed using NIH Toolbox (NIHTB) scores, including the List Sorting Working Memory Task, Flanker Inhibitory Control Task, and Pattern Comparison Processing Speed Task[36]. Uncorrected standard scores from the NIHTB assessments were used per recommendation by the ABCD neurocognitive working group[37].

Finally, a cumulative sum ACE index was constructed using binary endorsement items from the KSADS-COMP screening module, Family History Assessment, parent-reported demographics survey, youth-reported parental monitoring survey, family environment scale, and the acceptance subscale from the children's report of parental behavioral inventory[38]. This previously-used composite index captures lifetime exposure to abuse, neglect, household dysfunction, family conflict, and related adversities as defined by the seminal ACE questionnaire[38]. More extensive details regarding derivations, psychometric properties, and the reliability of measures described in the current study can be found in Barch et al. (2018)[26].

Figure 1: **Early Life Risk Factors for BD Used in Baseline Clustering**

Outcome Measures and Variables

KSADS-COMP–derived BD diagnoses were constructed from combined past and present parent and youth reports collected at baseline, 2-, 4-, and 6-year follow-up assessments. At each timepoint, parent and youth reports were combined to create binary (i.e., presence or absence) indicators of lifetime occurrence of bipolar I disorder (BD-I), bipolar II disorder (BD-II), other specified bipolar disorder (BD-NOS), and any bipolar spectrum disorder (BSD; incorporating any BD-I, BD-II, or other specified bipolar disorder endorsements) as of each assessment timepoint. KSADS-COMP diagnoses were generated in alignment with recommendations by Barch et al. (2021)[39]; see Supplementary Methods for details.

Data Pre-Processing

All analyses were conducted in R version 4.4.2[40]. Participants with missing or invalid values on any baseline BD risk variable were excluded from clustering and further analyses (n = 2,675 of 11,868 baseline participants; i.e., 22.5% exclusion). Statistical analyses of BD emergence were also conducted among clustered participants with complete cluster, covariate, and KSADS-COMP data at each relevant follow-up timepoint.

Clustering Procedure

To identify BD risk variable clusters at baseline, we employed the k-prototypes clustering algorithm, an unsupervised learning method suitable for mixed-type numeric and categorical data[41], via the *clustMixType* package (version 0.4-2)[42]. Briefly, this algorithm partitions data into clusters using a hybrid dissimilarity measure calculated from distances between data points and each cluster centroid for numeric variables, and matching dissimilarity for categorical variables, balanced by a weighting parameter lambda (λ).

Importantly, since clustering relies on distances, we z-scored numerical features to allow comparison between features with different ranges[43]. Additionally, we used the default data-driven λ provided by *clustMixType*, given no clear rationale for differential variable weighting, and no improvement in clustering performance following extensive λ-tuning analyses. See the Supplementary Methods for details.

To identify the empirically optimal number of clusters (*k*), candidate solutions *k* = 2:8 were assessed using internal validation. Specifically, the mean silhouette score (derived via the *validation_kproto* function from *clustMixType*[42]) across participant datapoints and clusters was utilized given its robust internal validation capabilities and efficient run-time[44]. The empirically optimal *k* was selected based on which candidate solution yielded the highest mean silhouette score.

The stability of the empirically optimal *k* solution was then evaluated via nonparametric bootstrap resampling of n = 1,000 samples, followed by re-clustering each data resample with the chosen z scaling method, λ, and *k*. Stability relative to the original partition was quantified using the Jaccard similarity coefficient, calculated via the *stability_kproto* function from *clustMixType*[42].

Sensitivity Analyses of the Clustering Procedure

Multiple sensitivity analyses were conducted to assess the robustness of the selected clustering solution. First, given that 22.5% of participants exhibited missingness at baseline, we conducted sensitivity analyses to examine whether the selected clustering solution was robust to missing-data handling. As an additional sensitivity analysis, we then examined whether the full 17-feature clustering solution could be reproduced with a smaller subset of baseline risk variables. Finally, we conducted analyses to determine whether the selected clustering solution depended on the k-prototypes algorithm. See the Supplementary Methods for details pertaining to each analysis.

Statistical Analyses

Model assumptions and associated diagnostics, including collinearity, autocorrelation, and residual normality, were evaluated using the *performance* package[45]. For discrete-time hazard models, primary inference was based on small-sample corrected CR2 robust standard errors clustered by participant[46]. Generalized estimating equation models used robust sandwich standard errors.

Onset of Bipolar Disorders by Risk Cluster

To examine when BD first emerged, we fit pooled logistic discrete-time hazard models on person-period data representing the intervals from baseline to 2-year follow-up, 2-year to 4-year follow-up, and 4-year to 6-year follow-up. Separate models were fit for BD-I, BD-II, BD-NOS, and any BSD. For each BD subtype, participants endorsing that subtype at baseline were excluded from the corresponding dataset. Intervals

after first onset were removed, and participants were otherwise right-censored at their last observed assessment.

Primary fixed effects included BD risk cluster, assessment timepoint indicators encoding the baseline hazard, sex-assigned-at-birth, and grand-mean centered between-person age[47]; models were estimated using pooled logistic regression with participant-clustered CR2 robust standard errors[46]. See the Supplementary Methods for additional covariate and model selection considerations.

After model fitting, we summarized model-based interval hazards, defined as the conditional probability of first onset within each interval given no prior onset, and the corresponding cumulative incidence by cluster across representative ages. For each cluster, predicted interval hazards were obtained by transforming the fitted linear predictor (via inverse-logit) at the median midpoint age for each interval, with other covariates held at reference values. Cumulative incidence was then derived using a life-table approach, in which survival was initialized at 1.00 at baseline, and updated after each interval by multiplying (1 - the predicted hazard) for that interval.

Occurrence of Bipolar Disorders Over Time by Risk Cluster

To estimate timepoint-specific occurrence of BD by cluster, we modeled timepoint-specific endorsement of BD-I, BD-II, BD-NOS, and any BSD using generalized estimating equations (GEE) with robust sandwich standard errors and an exchangeable working correlation via the *geepack* package[48]. See Supplementary Methods for more information related to the selected correlation structure. Fixed effects included BD risk cluster, assessment timepoint, the cluster by timepoint interaction, grand-mean–centered age at each timepoint, baseline status of the focal BSD outcome, and sex-assigned-at-birth. Thus, the BD-I model adjusted for baseline BD-I status, the BD-II model for baseline BD-II status, the BD-NOS model for baseline BD-NOS status, and the any BSD model for baseline any BSD status. This covariate was included so that cluster contrasts reflect prospective differences in occurrence of the focal BSD outcome beyond diagnoses already present at the time of cluster formation. Baseline BSD status was included so that cluster contrasts reflect prospective differences in BSD occurrence beyond diagnoses present at the time of cluster formation. Using fitted GEE models, we estimated timepoint-specific odds ratios comparing risk clusters from the cluster-by-timepoint interaction using the *emmeans* package[49].

Common Comorbidity Sensitivity Analysis

As a sensitivity analysis, we examined whether risk-cluster differences extended beyond BSDs to attention-deficit/hyperactivity disorder, anxiety disorders, depressive disorders, post-traumatic stress disorder, oppositional defiance disorder, and conduct disorder. See the Supplemental Methods for details.

Utility of Alternative Baseline Risk Representations Sensitivity Analysis

Lastly, we examined how much downstream BSD-related information was retained when the full multidomain baseline risk structure was represented in 3 alternative ways within the primary BD models: the 2-cluster label, the minimally viable feature subset (MVFS), and the full 17-variable baseline risk-variable set. See the Supplementary Methods for details.

Results

Sample and Clusters

Of 11,868 ABCD release 6.1 baseline participants, 77.5% (n = 9,193) had complete data on all 17 baseline BD risk variables and were included in clustering (Table 2). Using the default λ = 3.32, internal validation favored *k* = 2 clusters (mean silhouette, 0.52), exceeding alternative solutions by 0.29 to 0.40 mean units. The selected 2-cluster solution yielded cluster 1 (n = 1,391) and cluster 2 (n = 7,802), hereafter designated as the higher-risk and lower-risk clusters, respectively. The k = 2 solution was also highly stable in bootstrap resampling (mean Jaccard = 0.98; SD = 0.01).

Table 2
Baseline Characteristics of the Clustered Sample

Variable	Cluster 1 (Higher-Risk Cluster)	Cluster 2 (Lower-Risk Cluster)	Total
N	1,391	7,802	9,193
Age in Years, Mean ± SD	9.92 ± 0.62	9.97 ± 0.62	9.96 ± 0.62
Sex-assigned-at-birth, n (%)			
Male	833 (59.9%)	4,015 (51.5%)	4,848 (52.7%)
Female	558 (40.1%)	3,787 (48.5%)	4,345 (47.3%)
BD-I Diagnosis, n (%)	37 (2.7%)	33 (0.4%)	70 (0.8%)
BD-II Diagnosis, n (%)	16 (1.2%)	8 (0.1%)	24 (0.3%)
BD-NOS Diagnosis, n (%)	50 (3.6%)	54 (0.7%)	104 (1.1%)
Any BSD Diagnosis, n (%)	83 (6.0%)	84 (1.1%)	167 (1.8%)

Note: BD-I = bipolar I disorder; BD-II = bipolar II disorder; BD-NOS = other specified bipolar disorder; BSD = bipolar spectrum disorder.

Cluster Profiles at Baseline

The higher- and lower-risk clusters differed significantly across all 17 baseline risk variables (all p<.001; Table 3). Relative to the lower-risk cluster, the higher-risk cluster showed greater CBCL DSM-5 depression and anxiety symptoms, attention problems, and aggression problem T-scores; higher sleep disturbance scale for children scores; greater lifetime GBI mania symptoms; and greater ACE exposure.

The higher-risk cluster also showed higher negative and positive urgency, more frequent family history of depression and mania, and more frequent bullying victimization. Neighborhood safety and Child Opportunity Index values were lower in the higher-risk cluster, and neurocognitive performance was lower across NIHTB Flanker, List Sorting, and Pattern Comparison measures.

Table 3
Baseline Risk Variable Descriptive Statistics by Cluster

Variable	Lower-Risk Cluster Mean ± SD | n (%)	Higher-Risk Cluster Mean ± SD | n (%)	Effect Size [95% CI]	P-Value
Numeric Variables				
ACE Index Sum Score	1.80 ± 1.28	2.85 ± 1.39	d = 0.81 [0.75, 0.87]	< .001
CBCL Aggression Problems T	51.19 ± 2.61	60.88 ± 8.17	d = 2.43 [2.36, 2.50]	< .001
CBCL Attention Problems T	52.13 ± 3.51	62.56 ± 8.67	d = 2.23 [2.17, 2.30]	< .001
CBCL DSM-5 Anxiety T	51.86 ± 3.49	61.62 ± 9.22	d = 2.03 [1.96, 2.09]	< .001
CBCL DSM-5 Depression T	51.87 ± 3.31	62.28 ± 7.32	d = 2.49 [2.43, 2.56]	< .001
COI Z-Score	0.01 ± 0.03	-0.001 ± 0.04	d = -0.39 [-0.44, -0.33]	< .001
GBI Mania Sum Score	0.55 ± 1.24	4.94 ± 4.62	d = 2.06 [1.99, 2.12]	< .001
Mean Neighborhood Safety	3.99 ± 0.91	3.56 ± 1.08	d = -0.45 [-0.51, -0.39]	< .001
NIHTB Flanker Score	94.57 ± 8.76	92.06 ± 10.12	d = -0.28 [-0.34, -0.22]	< .001
NIHTB List Sorting Score	97.75 ± 11.71	93.09 ± 12.49	d = -0.39 [-0.45, -0.34]	< .001
NIHTB Pattern Comparison Score	88.85 ± 14.36	85.62 ± 14.87	d = -0.22 [-0.28, -0.17]	< .001
SDSC Sum Score	34.53 ± 5.80	46.30 ± 11.05	d = 1.72 [1.66, 1.78]	< .001
UPPS-P Negative Urgency Score	8.30 ± 2.55	9.50 ± 2.83	d = 0.46 [0.40, 0.52]	< .001
UPPS-P Positive Urgency Score	7.76 ± 2.85	8.96 ± 3.18	d = 0.41 [0.36, 0.47]	< .001

Note: For categorical variables, n and percent of positive endorsement are shown. Abbreviations: ACE=adverse childhood experience; CI=confidence interval; COI=child opportunity index; CBCL=child behavior checklist; GBI=general behavior inventory; NIHTB = NIH toolbox; SD=standard deviation; SDSC=sleep disturbance scale for children; UPPS-P=urgency, premeditation, perseverance, sensation seeking, and positive urgency impulsive behavior scale for children.

Variable	Lower-Risk Cluster Mean ± SD | n (%)	Higher-Risk Cluster Mean ± SD | n (%)	Effect Size [95% CI]	P-Value
Categorical Variables				
Bullying Victimization	1548 (19.8%)	786 (56.5%)	OR = 5.25 [4.66, 5.92]	< .001
Family History of Depression	1981 (25.4%)	854 (61.4%)	OR = 4.67 [4.15, 5.27]	< .001
Family History of Mania	297 (3.8%)	185 (13.3%)	OR = 3.88 [3.19, 4.70]	< .001

Note: For categorical variables, n and percent of positive endorsement are shown. Abbreviations: ACE=adverse childhood experience; CI=confidence interval; COI=child opportunity index; CBCL=child behavior checklist; GBI=general behavior inventory; NIHTB = NIH toolbox; SD=standard deviation; SDSC=sleep disturbance scale for children; UPPS-P=urgency, premeditation, perseverance, sensation seeking, and positive urgency impulsive behavior scale for children.

Missing-Data and Imputation Sensitivity Analysis

In the analysis of missing data, baseline missingness was modest overall, although missingness was not consistent with missing completely at random (Little's χ^2=1308.98, df = 767, p<.001). At the baseline assessment timepoint, 11,865 of 11,868 participants (99.97%; Table S3) were eligible for imputation. Refit clustering of the imputed data again favored *k* = 2 (mean silhouette = 0.509; Table S4), and the imputed *k* = 2 solution showed strong agreement with the original complete-case anchor partition (adjusted Rand index = 0.871; Table S4) and high stability (mean Jaccard = 0.938; Table S4). Full details are provided in the Supplementary Results.

Feature Importance and Parsimonious Feature-Set Sensitivity Analysis

In the feature-importance and parsimony sensitivity analysis, the strongest and most consistent contributors across analyses were parent-reported CBCL depression, aggression, attention, and anxiety T-scores; followed by lifetime GBI mania symptoms; etc. See Table S5 for full results. In forward wrapper selection, the MVFS meeting reproduction and stability criteria was comprised of 7 features: parent-reported CBCL DSM-oriented depression, DSM-oriented anxiety, aggression, and attention T-scores; lifetime GBI mania symptoms; sleep disturbance; and family history of depression. Details are provided in the Supplementary Results.

Algorithm Benchmark Sensitivity Analysis

Regarding potential specificity of the 2-cluster solution to k-prototypes, the 2-cluster solution remained preferred across algorithms. More specifically, the highest mean silhouette values for each algorithm were observed at *k* = 2 (Table S6). See the Supplementary Results for details.

Onset of Bipolar Disorders by Risk Cluster

A detailed characterization of this analytical subsample is provided in Table S1. All DTH models used participant-clustered CR2 robust standard errors. Interval probabilities of first onset were highest from baseline to year-2 follow-up for BD-I and any BSD and were consistently greater in the higher-risk cluster, indicating earlier concentration of onsets (Fig. 2A-D). Overall, 598 of 9,193 clustered participants (6.5%) endorsed any BSD by year-6 follow-up. Across intervals, membership in the higher-risk cluster was associated with significantly greater odds of first onset for BD-I (OR = 2.99, 95% CI [2.20, 4.06], p<.001), BD-II (OR = 2.27, 95% CI [1.35, 3.80], p = 0.002), BD-NOS (OR = 3.36, 95% CI [2.51, 4.51], p<.001), and any BSD (OR = 2.74, 95% CI [2.21, 3.41], p<.001).

Cumulative incidence diverged steadily through year-6 follow-up, with greater model-based cumulative first-onset risk in the higher-risk cluster across all BSDs (Fig. 2E-H). At the timepoint-specific median ages and with other covariates held at reference values, cumulative first-onset risk in the higher-risk vs lower-risk cluster was 9.6% vs 3.3% for BD-I (risk difference [RD] = 6.3 percentage points, 95% CI [3.7, 8.7], p<.001; risk ratio [RR] = 2.86, 95% CI [2.14, 3.82]), 3.6% vs 1.6% for BD-II (RD = 2.0 percentage points, 95% CI [0.3, 3.6], p = 0.002; RR = 2.24, 95% CI [1.34, 3.73]), 11.5% vs 3.6% for BD-NOS (RD = 7.9 percentage points, 95% CI [5.1, 10.7], p<.001; RR = 3.17, 95% CI [2.40, 4.19]), and 19.7% vs 7.9% for any BSD (RD = 11.8 percentage points, 95% CI [8.4, 15.2], p<.001; RR = 2.50, 95% CI [2.06, 3.03]).

Figure 2: **Discrete Time Hazards and Cumulative Incidence of Bipolar Disorder Onset by Risk Cluster**

Occurrence of Bipolar Disorders Over Time by Risk Cluster

A detailed characterization of this analytical subsample is provided in Table S2. The higher-risk cluster showed significantly greater odds of occurrence at year-2 follow-up for BD-I (OR = 6.38, 95% CI [3.78, 10.79], p<.001), BD-II (OR = 11.79, 95% CI [5.01, 27.72], p<.001), BD-NOS (OR = 5.86, 95% CI [3.57, 9.63], p<.001), and any BSD (OR = 6.48, 95% CI [4.43, 9.49], p<.001), controlling for timepoint, age, baseline BD status, and sex-assigned-at-birth.

At year-4 follow-up, odds remained significantly greater in the higher-risk cluster for BD-I (OR = 2.00, 95% CI [1.25, 3.22], p = 0.004), BD-NOS (OR = 2.66, 95% CI [1.72, 4.11], p<.001), and any BSD (OR = 1.97, 95% CI [1.44, 2.69], p<.001), but not BD-II (OR = 0.97, 95% CI [0.38, 2.53], p=.957). At year-6 follow-up, odds remained greater in the higher-risk cluster for BD-I (OR = 2.74, 95% CI [1.56, 4.82], p<.001), BD-NOS (OR = 3.55, 95% CI [2.12, 5.97], p<.001), and any BSD (OR = 2.45, 95% CI [1.67, 3.58], p<.001), but not BD-II (OR = 0.67, 95% CI [0.16, 2.89], p = 0.594).

Common Comorbidity Sensitivity Analysis

In the sensitivity analysis examining relationships between risk clusters and conditions commonly comorbid with BD, the higher-risk cluster showed greater odds of broader non-BD psychopathology over follow-ups than the lower-risk cluster, but effects were larger for BSD occurring alongside a concurrent comorbid diagnosis. Additionally, in direct pairwise contrasts, higher-risk cluster membership more

strongly differentiated BSD and a concurrent comorbidity than other diagnoses only at year-2. See the Supplementary Results for details.

Utility of Alternative Baseline Risk Representations Sensitivity Analysis

In the sensitivity analysis comparing alternative baseline risk representations, the full 17-variable baseline risk set retained the most downstream BSD-related information, as reflected by the highest McFadden pseudo-R^2, whereas the 7-variable MVFS was generally intermediate between the full model and the empirically derived 2-cluster label. Although fuller baseline risk representations showed somewhat better model fit overall, between-model differences in McFadden pseudo-R^2, Brier score, and log-loss were modest, and the cluster-label effects in the primary onset and occurrence models remained substantially larger than the effects of any individual baseline risk variable. Full details are provided in the Supplementary Results.

Discussion

In this large, demographically diverse cohort followed from late childhood into mid-adolescence, multidomain childhood BD risk factors were organized into a stable, clinically interpretable 2-cluster structure. Using 17 literature-grounded, staging-aligned baseline features, we found that risk in the general population separated into a higher- or lower-risk profile without additional intermediary subgroups. The higher-risk cluster, characterized by greater clinical/prodromal symptoms, impulsivity, sleep disturbance, family history of mood disorders, adverse experiences, environmental disadvantage, and lower neurocognitive performance, showed elevated first onset and occurrence of BSDs through age 16. By year-6 follow-up, youths in the higher-risk cluster had roughly threefold greater cumulative risk of developing any BSD than youth in the lower-risk cluster. Onsets were also not uniformly distributed across follow-ups: model-based interval hazards suggested earlier concentration of risk in the higher-risk cluster, with BD-I and any BSD hazards highest from baseline to year-2 follow-up.

The 2-cluster solution is notable not only because it was associated with later BSD emergence, but because it appears to capture the most separable and stable profile of known childhood BD risk factors in this developmental sample. In other words, rather than testing how individual risk factors predict later BSD, the present study used data-driven clustering to quantify how known risk factors naturally organized in childhood independent of downstream BD information. Many early-life indicators associated with BD are individually transdiagnostic, and either modestly discriminative or lacking complete prognostic utility on their own. Accordingly, defining how risk co-occurs across domains may be an important step toward clarifying which combinations of factors signal developmentally meaningful risk.

These findings extend prior longitudinal work in several ways. First, they replicate findings from familial high-risk cohorts showing that subsyndromal mood symptoms, impulsivity, sleep disturbance, and family

history precede BSD onset,[21,23] but do so in a population-based sample with broader multidomain risk representation. Second, they complement person-level risk calculators that combine familial and early symptomatic indicators[23] by characterizing how multidomain BD risk factors co-occur in childhood before syndromal illness emerges. More broadly, these findings align with developmental staging models and provide an empirically grounded framework for risk stratification that can be further tested, validated, and clinically translated.

The identified higher-risk cluster may have important implications for developmental risk stratification. The component measures underlying this profile, particularly that of the MVFS, are relatively accessible and include questionnaires and family history that can be obtained in routine clinical settings. While the current findings do not support clinical deployment as a diagnostic screening tool, they demonstrate meaningful risk enrichment. By year-6 follow-up, 6.5% of clustered youth had developed any BSD, including 14.7% in the higher-risk cluster and 5.0% in the lower-risk cluster. Although not directly comparable, this magnitude of risk enrichment is notable given that prior work in a familial high-risk sample reported an overall 5-year BSD conversion rate of 13.1%[23]. Accordingly, the most appropriate implication of the current findings is not deterministic prediction, but that these clusters enable targeted risk stratification. For example, given the early concentration of onset risk particularly across the transition from late childhood to early adolescence, youth in the higher-risk cluster may benefit from developmental symptom monitoring, psychoeducation about early warning signs, sleep hygiene and routine stabilization, and other low-burden preventive strategies.

The present findings are strengthened by converging evidence across sensitivity analyses. First, the 2-cluster profile remained robust after missing data imputation. Second, much of the full 17-feature solution could be reproduced with a 7-feature MVFS comprised of parent-reported mood, anxiety, behavioral, sleep, and familial-loading indicators, suggesting the broader profile was not contingent on the full set of variables while still representing their distributions. Third, the 2-cluster structure was reliably supported across alternative clustering algorithms. Finally, the common-comorbidity sensitivity analysis revealed that higher-risk cluster membership was associated with elevated psychopathology broadly, not only BSD, though effects were strongest for BSD occurring alongside concurrent comorbidity(s). This partial specificity is consistent with the known overlap among early risk indicators across disorders and aligns with findings that comorbidity is common in adolescent BD[50]. Together, these results indicate the identified clusters capture a robust developmental psychopathology risk profile enriched for BSDs, rather than BD-exclusive risk.

These implications should also be considered from the perspective of health equity. Several variables contributing to the multidomain risk profile, including child opportunity index and perceived neighborhood safety, reflect social and environmental conditions that are unequally distributed across children and their families. Accordingly, future use of risk stratification approaches should avoid treating risk profiles as fixed individual liabilities detached from their structural context. Furthermore, future validation should evaluate whether risk profiles are similarly calibrated across underrepresented and

minoritized populations, and whether preventive strategies informed by these profiles reduce rather than reinforce disparities in access to early identification and care.

Key strengths of this study include the use of a prespecified, theory-driven set of childhood BD risk indicators, the application of a mixed-type clustering framework suited to those data, and prospective follow-up across multiple adolescent timepoints. Importantly, rather than examining isolated predictors or strictly enriched familial cohorts, the study quantified multidomain BD risk profiles in a population-based sample and then tested whether that structure was prospectively associated with both first onset and occurrence of BSD. The large ABCD cohort also enabled calibration of risk under population base rates, reducing referral and familial-loading biases inherent to clinical high-risk studies. Finally, the convergent pattern across onset, occurrence, imputation, parsimony, and algorithm-benchmark analyses supports the view that the identified 2-cluster solution is both interpretable and reasonably robust.

Despite these strengths, several limitations warrant consideration. First, the primary clustering pipeline was conducted in the complete-case baseline sample. Although imputation-based sensitivity analyses suggested that the 2-cluster profile was robust to baseline missing-data handling, complete-case restriction may still have affected the generalizability of the cluster distributions. Second, BD diagnoses were derived from the KSADS-COMP, rather than clinician-administered diagnostic interviews. Although the KSADS-COMP offers efficiency and broad coverage in large cohorts, prior work has noted moderate concordance with clinician-rated diagnoses, and the v1.0 algorithms used at ABCD baseline and year-2 follow-up appeared to overestimate past manic episodes[35,39]. We applied KSADS-COMP scoring algorithms recommended by Barch[39] to mitigate false positives, but some diagnostic misclassification remains possible. BD prevalence rates reported in the current study should therefore not be interpreted as population estimates. Third, despite the large overall cohort, data at year-6 follow-up was only available for a subset of participants, and BSD remained relatively uncommon by mid-adolescence (particularly for BD-II), resulting in wide confidence intervals for some model estimates. The exact magnitude of effect sizes for BD-II should therefore be interpreted cautiously. Finally, the observation window extended only through year-6 follow-up, and therefore captures early adolescent BSD emergence. Additional follow-up is needed to determine whether the identified risk profile continues to stratify BSD incidence into later adolescence and early adulthood, or whether its utility fluctuates as developmental context and risk exposures evolve.

Given the above strengths and limitations, several avenues for future research emerge. First, the broad psychopathology associations and partial BSD plus concurrent comorbidity specificity pattern observed here warrants more refined work on diagnostic specificity. Specifically, future studies should examine whether incorporating additional disorder-specific markers or prediction-oriented approaches yield more differentiated developmental profiles across BD, depressive, anxiety, and externalizing symptoms. Second, longer follow-up will be necessary to determine whether the present 2-cluster solution primarily captures risk for earlier BSD emergence, or whether it continues to stratify incidence later in development. Third, external validation in cohorts with clinician-rated diagnoses, biological measures, and more intensive prospective phenotyping will be important for clarifying which youths within the

higher-risk cluster ultimately convert to BSDs, and why. Finally, future work should examine whether this multidomain childhood risk profile can be integrated with other stratification tools, including clinically ascertained family history of BD, polygenic risk, and longitudinal symptom trajectories, to yield more developmentally precise and clinically actionable models of BD risk.

In summary, we identified a stable and clinically interpretable 2-cluster profile of childhood BD risk at ages 9–10 spanning clinical/prodromal symptoms, familial loading, environmental factors, and neurocognitive domains. This multidomain risk structure was prospectively associated with both the timing and occurrence of adolescent BSD emergence and remained robust across missing-data, feature-reduction, and cross-algorithm sensitivity analyses. The higher-risk cluster also indexed broad psychopathology over time, while showing especially strong early enrichment for BSD occurring alongside concurrent comorbidities. Together, these findings suggest that multidomain childhood BD risk factors may coalesce into a robust developmental high-risk profile that can be quantified before syndromal illness emerges. If replicated and externally validated, this 2-cluster solution could provide an empirical foundation for early-stage bipolar risk stratification, helping identify the subset of youth who may benefit most from closer monitoring and preventive intervention during the transition into adolescence.

Declarations

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development™ (ABCD) Study, held in the NIH Brain Development Cohorts Data Sharing Platform. This is a multisite, longitudinal study designed to recruit more than 10,000children aged 9–10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers: U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018,U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117,U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156,U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093,U01DA041089, U24DA041123, U24DA041147. Additional support for this work was made possible from NIEHS R01-ES032295 and R01-ES031074. A full list of supporters is available at Federal Partners – ABCD Study (https://abcdstudy.org/about/federal-partners/). ABCD Consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD Consortium investigators. The ABCD data repository grows

and changes over time. The ABCD data used in this report came from (https://doi.org/10.82525/jy7n-g441). The research reported in this publication used computational infrastructure supported by the Office of Research Infrastructure Programs, Office of the Director, of the National Institutes of Health under Award Number S10OD034224. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

We would like to thank the children and families who have and continue to participate in the ABCD Study. We would also like to thank Dr. Kenneth Koback, PhD, for valuable insight into the diagnostic algorithms underlying the KSADS-COMP platform, and advice on how resulting data from them should be utilized and interpreted. Finally, we would like to thank Drs. Mary L. Phillips, MD; Martin Paulus, MD; and Ran Barzilay, MD-PhD, for engaging feedback provided at the 2025 Society of Biological Psychiatry annual meeting, which aided in improving the analytical framework of this study.

References

1. Van Meter, A. R., Moreira, A. L., & Youngstrom, E. A. (2011). Meta-analysis of epidemiologic studies of pediatric bipolar disorder. *J Clin Psychiatry*, *72*(9), 1250–1256.
2. Scott, J., Graham, A., Yung, A., Morgan, C., Bellivier, F., & Etain, B. (2022). A systematic review and meta-analysis of delayed help-seeking, delayed diagnosis and duration of untreated illness in bipolar disorders. *Acta Psychiatr Scand*, *146*(5), 389–405.
3. Goldstein, T. R., Birmaher, B., Axelson, D., Goldstein, B. I., Gill, M. K., Esposito-Smythers, C. *et al*. (2009). Psychosocial functioning among bipolar youth. *J Affect Disord*, *114*(1–3), 174–183.
4. Mena, S. A., Verdolini, N., Miret, S., Laherrán, N., Pàmpols, S., & Palacios-Garrán, R. (2025). Impact of Duration of Untreated Bipolar Disorder on Clinical Outcomes: A Systematic Review. *Early intervention in psychiatry*, *19*(6), e70059.
5. Cirone, C., Secci, I., Favole, I., Ricci, F., Amianto, F., Davico, C. *et al*. (2021). What Do We Know about the Long-Term Course of Early Onset Bipolar Disorder? A Review of the Current Evidence. *Brain Sci*, *11*(3), 341.
6. Hafeman, D. M., Merranko, J., Axelson, D., Goldstein, B. I., Goldstein, T., Monk, K. *et al*. (2016). Toward the Definition of a Bipolar Prodrome: Dimensional Predictors of Bipolar Spectrum Disorders in At-Risk Youths. *Am J Psychiatry*, *173*(7), 695–704.
7. Brancati, G. E., Perugi, G., Milone, A., Masi, G., & Sesso, G. (2021). Development of bipolar disorder in patients with attention-deficit/hyperactivity disorder: A systematic review and meta-analysis of prospective studies. *J Affect Disord*, *293*, 186–196.
8. Aymerich, C., Bullock, E., Rowe, S. M. B., Catalan, A., & Salazar de Pablo, G. (2025). Aggressive Behavior in Children and Adolescents With Bipolar Spectrum Disorder: A Systematic Review of the Prevalence, Associated Factors, and Treatment. *JAACAP Open*, *3*(1), 42–55.
9. Salazar de Pablo, G., Cabras, A., Pereira, J., Castro Santos, H., de Diego, H., Catalan, A. *et al*. (2023). Predicting bipolar disorder I/II in individuals at clinical high-risk: Results from a systematic review. *J*

Affect Disord, *325*, 778–786.

10. Johnson, S. L., Carver, C. S., & Tharp, J. A. (2017). Suicidality in Bipolar Disorder: The Role of Emotion-Triggered Impulsivity. *Suicide Life Threat Behav*, *47*(2), 177–192.
11. Wilde, A., Chan, H. N., Rahman, B., Meiser, B., Mitchell, P. B., Schofield, P. R. *et al*. (2014). A meta-analysis of the risk of major affective disorder in relatives of individuals affected by major depressive disorder or bipolar disorder. *J Affect Disord*, *158*, 37–47.
12. O'Connell, K. S., & Coombes, B. J. (2021). Genetic contributions to bipolar disorder: current status and future directions. *Psychol Med*, *51*(13), 2156–2167.
13. Leboyer, M., & Kupfer, D. J. (2010). Bipolar disorder: new perspectives in health care and prevention. *J Clin Psychiatry*, *71*(12), 1689–1695.
14. Baranyi, G., Di Marco, M. H., Russ, T. C., Dibben, C., & Pearce, J. (2021). The impact of neighbourhood crime on mental health: A systematic review and meta-analysis. *Soc Sci Med*, *282*, 114106.
15. Manoli, A., Wright, L. C., Shakoor, S., Fisher, H. L., & Hosang, G. M. (2023). The association between childhood bullying victimisation and childhood maltreatment with the clinical expression of bipolar disorder. *J Psychiatr Res*, *158*, 226–230.
16. Harvey, A. G., Talbot, L. S., & Gershon, A. (2009). Sleep Disturbance in Bipolar Disorder Across the Lifespan. *Clin Psychol (New York)*, *16*(2), 256–277.
17. Keramatian, K., Torres, I. J., & Yatham, L. N. (2022). Neurocognitive functioning in bipolar disorder: What we know and what we don't. *Dialogues Clin Neurosci*, *23*(1), 29–38.
18. Lombardo, L. E., Bearden, C. E., Barrett, J., Brumbaugh, M. S., Pittman, B., Frangou, S. *et al*. (2012). Trait impulsivity as an endophenotype for bipolar I disorder. *Bipolar Disord*, *14*(5), 565–570.
19. Köhler-Forsberg, O., Ge, F., Hauksdóttir, A., Thordardottir, E. B., Ásbjörnsdóttir, K., Rúnarsdóttir, H., Tómasson, G. *et al*. (2024). Adverse childhood experiences and psychological functioning among women with schizophrenia or bipolar disorder: population-based study. *Br J Psychiatry*, *224*(1), 6–12.
20. Park, YM., Shekhtman, T. & Kelsoe, J.R. (2020). Interaction between adverse childhood experiences and polygenic risk in patients with bipolar disorder. *Transl Psychiatry*, *10*, 326.
21. Birmaher, B., Merranko, J., Hafeman, D., Goldstein, B. I., Diler, R., Levenson, J. C. *et al*. (2021). A Longitudinal Study of Psychiatric Disorders in Offspring of Parents With Bipolar Disorder From Preschool to Adolescence. *J Am Acad Child Adolesc Psychiatry*, *60*(11), 1419–1429.
22. Birmaher, B., Axelson, D., Goldstein, B., Strober, M., Gill, M. K., Hunt, J. *et al*. (2009). Four-year longitudinal course of children and adolescents with bipolar spectrum disorders: the Course and Outcome of Bipolar Youth (COBY) study. *Am J Psychiatry*, *166*(7), 795–804.
23. Hafeman, D. M., Merranko, J., Goldstein, T. R., Axelson, D., Goldstein, B. I., Monk, K. *et al*. (2017). Assessment of a Person-Level Risk Calculator to Predict New-Onset Bipolar Spectrum Disorder in Youth at Familial Risk. *JAMA psychiatry*, *74*(8), 841–847.

24. Goldstein, B. I., Birmaher, B., Carlson, G. A., DelBello, M. P., Findling, R. L., Fristad, M. *et al*. (2017). The International Society for Bipolar Disorders Task Force report on pediatric bipolar disorder: Knowledge to date and directions for future research. *Bipolar Disord*, *19*(7), 524–543.
25. Duffy A. (2014). Toward a comprehensive clinical staging model for bipolar disorder: integrating the evidence. *Can J Psychiatry*, *59*(12), 659–666.
26. Barch, D. M., Albaugh, M. D., Avenevoli, S., Chang, L., Clark, D. B., Glantz, M. D., et al. (2018). Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. *Dev Cogn Neurosci*, *32*, 55–66.
27. Garavan, H., H. Bartsch, K. Conway, A. Decastro, R.Z. Goldstein, S. Heeringa, T. *et al*. (2018). Recruiting the ABCD Sample: Design Considerations and Procedures. *Dev Cogn Neurosci*, *32*, 16–22.
28. Achenbach, T. M. (2001). Manual for ASEBA school-age forms & profiles. *University of Vermont*, Research Center for Children, Youth & Families.
29. Bruni, O., Ottaviano, S., Guidetti, V., Romoli, M., Innocenzi, M., Cortesi, F. *et al*. (1996). The Sleep Disturbance Scale for Children (SDSC). Construction and validation of an instrument to evaluate sleep disturbances in childhood and adolescence. *J Sleep Res*, *5*(4), 251–261.
30. Youngstrom, E. A., Perez Algorta, G., Youngstrom, J. K., Frazier, T. W., & Findling, R. L. (2021). Evaluating and Validating GBI Mania and Depression Short Forms for Self-Report of Mood Symptoms. *J Clin Child Adolesc Psychol*, *50*(5), 579–595.
31. Cyders M.A., Littlefield A.K., Coffey S., Karyadi K.A. (2014). Examination of a short English version of the UPPS-P Impulsive Behavior Scale. *Addict Behav. 39*(9):1372–6.
32. Rice, J. P., Reich, T., Bucholz, K. K., Neuman, R. J., Fishman, R., Rochberg, N. *et al*. (1995). Comparison of direct interview and family history diagnoses of alcohol dependence. *Alcohol Clin Exp Res*, *19*(4), 1018–1023.
33. Dolores, A.G., Noelke, C., McArdle, N., Sofer, N., Huntington, N., Hardy, E. *et al*. (2020). "The Geography of Child Opportunity: Why Neighborhoods Matter for Equity." http://diversitydatakids.org/research-library/research-report/geography-child-opportunity-why-neighborhoods-matter-equity
34. Hamilton, C. M., Strader, L. C., Pratt, J. G., Maiese, D., Hendershot, T., Kwok, R. K. *et al*. (2011). The PhenX Toolkit: get the most from your measures. *Am J Epidemiol*, *174*(3), 253–260.
35. Townsend, L., Kobak, K., Kearney, C., Milham, M., Andreotti, C., Escalera, J., *et al*. (2020). Development of Three Web-Based Computerized Versions of the Kiddie Schedule for Affective Disorders and Schizophrenia Child Psychiatric Diagnostic Interview: Preliminary Validity Data. *J Am Acad Child Adolesc Psychiatry*, *59*(2), 309–325.
36. Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J. *et al*. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(11 Suppl 3), S54–S64.
37. Adolescent Brain Cognitive Development Study (2025). *NIH Toolbox (Cognition)*. https://docs.abcdstudy.org/latest/documentation/non_imaging/nc.html#nc_y_nihtb

38. Breslin, F.J., Ratliff, E.L., Cohen, Z.P, Kroff, J.M., Kerr, K.L. (2025). Measuring adversity in the ABCD® Study: systematic review and recommendations for best practices. *BMC Med Res Methodol*, *25*, 77.
39. Barch, D. M., Albaugh, M. D., Baskin-Sommers, A., Bryant, B. E., Clark, D. B., Dick, A. S. *et al*. (2021). Demographic and mental health assessments in the adolescent brain and cognitive development study: Updates and age-related trajectories. *Dev Cogn Neurosci*, *52*, 101031.
40. R Core Team (2025). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL https://www.R-project.org/
41. Huang, Z. (1998). Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values. *Data Min Knowl Discov*, *2*, 283–304.
42. Szepannek G (2018). "clustMixType: User-Friendly Clustering of Mixed-Type Data in R." *The R Journal*, https://doi.org/10.32614/RJ-2018-048
43. Wongoutong C. (2024). The impact of neglecting feature scaling in k-means clustering. *PloS one*, *19*(12), e0310839. https://doi.org/10.1371/journal.pone.0310839
44. Todeschini, R., Davide, B., Termopoli, V., Consonni, V. (2024). Extended multivariate comparison of 68 cluster validity indices. A review. *Chemometrics and Intelligent Laboratory Systems*, *251*, 105117.
45. Lüdecke D., Ben-Shachar M., Patil I., Waggoner P., Makowski D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139.
46. Pustejovsky, J. E., & Tipton, E. (2017). Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models. *J Bus Econ Stat*, *36*(4), 672–683.
47. Hoffman, L., & Stawski, R.S. (2009). Persons as Contexts: Evaluating Between-Person and Within-Person Effects in Longitudinal Analysis. *Res Hum Dev*, *6*(2–3), 97–120.
48. Halekoh U., Højsgaard S., Yan J. (2005). The R Package geepack for Generalized Estimating Equations. *J Stat Softw*, *15*(2), 1–11.
49. Lenth R (2025). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version 1.11.2*-80001, https://rvlenth.github.io/emmeans/
50. Sesso, G., Brancati, G. E., Masi, G. (2023). Comorbidities in Youth with Bipolar Disorder: Clinical Features and Pharmacological Management. *Curr Neuropharmacol*, *21*(4), 911–934.

Figures

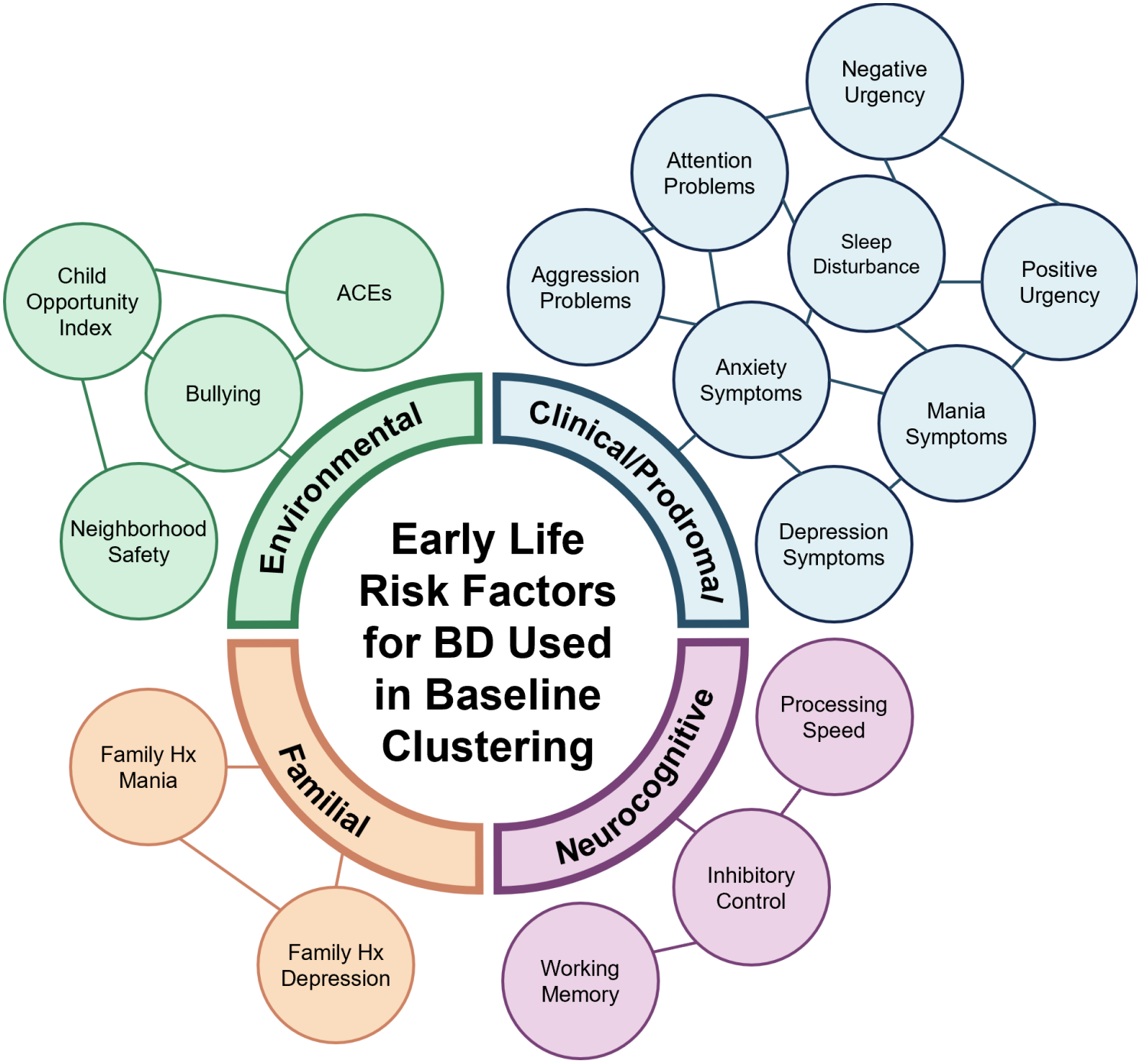


Figure 1

Early Life Risk Factors for BD Used in Baseline Clustering

Note: Individual risk factors associated with the manifestation of BD are color coded according to their corresponding domain. Abbreviations: ACEs = adverse childhood experiences; Hx = history.

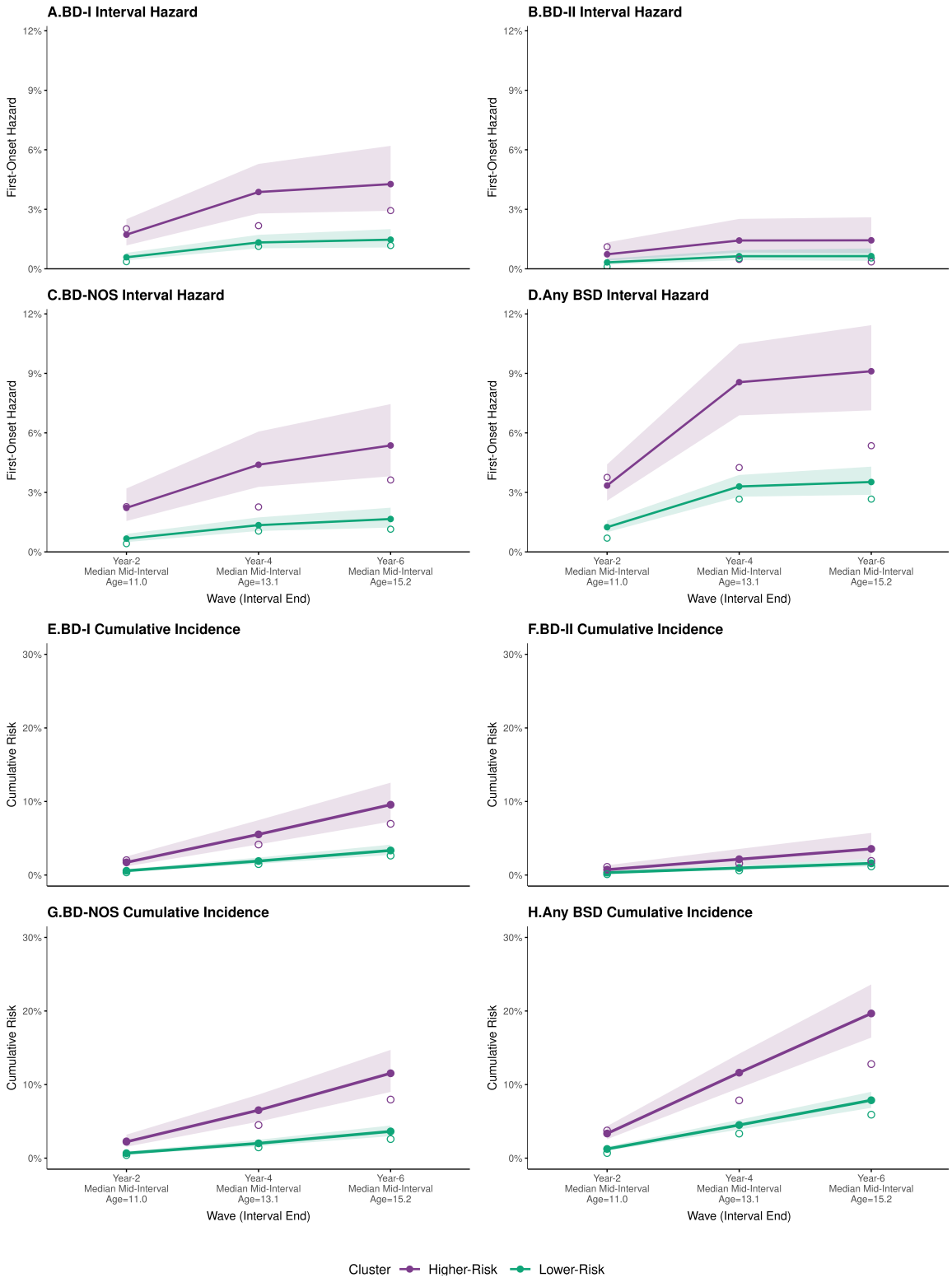


Figure 2

Discrete Time Hazards and Cumulative Incidence of Bipolar Disorder Onset by Risk Cluster

Note: (**A**) BD-I interval hazard; (**B**) BD- II interval hazard; (**C**) BD-NOS interval hazard; (**D**) Any BSD interval hazard; (**E**) BD-I cumulative incidence; (**F**) BD-II cumulative incidence; (**G**) BD-NOS cumulative incidence; (**H**) Any BSD cumulative incidence. Lines connect model-based predicted interval hazards or cumulative

incidence estimates of first onset at each follow-up timepoint (i.e., year-2, year-4, year-6). Shaded ribbons indicate pointwise 95% model-based CI around the predicted hazards or derived cumulative incidence estimates. Open circles show observed per-timepoint event proportions among those at risk. Observed proportions are unadjusted, whereas model-based estimates are adjusted predictions evaluated at the timepoint-specific median mid-interval age displayed under the x-axis tick, with covariates at reference levels. Abbreviations: BD-I = bipolar I disorder; BD-II = bipolar II disorder; BD-NOS = other specified bipolar disorder; BSD = bipolar spectrum disorder.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalInformation.docx
- SupplementalTables.xlsx