Examining the Most Important Risk Factors Predicting Persistent and Distressing Psychotic-like Experiences in Youth

Nicole R. Karcher, Aristeidis Sotiras, Tara A. Niendam, Elaine F. Walker, Joshua J. Jackson, and Deanna M. Barch

1 Washington University School of Medicine, Department of Psychiatry
2 Washington University School of Medicine, Department of Radiology
3 Washington University School of Medicine, Institute for Informatics, Data Science & Biostatistics
4 University of California, Davis, Dept. of Psychiatry and Behavioral Sciences
5 Emory University, Dept. of Psychology
6 Washington University in St. Louis, Department of Psychological and Brain Sciences

Please address all correspondence to: Nicole Karcher, Dept. of Psychiatry, Campus Box 1125, Washington University in St. Louis, One Brookings Drive, Saint Louis, MO 63110; Phone number: 314-935-8547; Email: nkarcher@wustl.edu

Abstract Word Count: 250
Text Body Word Count: 3,999
Abstract

**Background:** Persistence and distress distinguish more clinically significant psychotic-like experiences (PLEs) from those that are less likely to be associated with impairment and/or need for care. Identifying risk factors that differentiate clinically relevant PLEs early in development is important for improving our understanding of the etiopathogenesis of these experiences. Machine learning analyses examined the most important baseline factors distinguishing persistent distressing PLEs.

**Methods:** Using Adolescent Brain Cognitive Development Study PLEs data over three time points (ages 9-13), individuals with persistent distressing PLEs (n=303), transient distressing PLEs (n=374), and demographically matched low-level PLEs groups were created. Random forest classification models were trained to distinguish among persistent distressing vs. low-level PLEs, transient distressing vs. low-level PLEs, and persistent distressing vs. transient distressing PLEs. Models were trained using identified baseline predictors as input features (i.e., cognitive, neural [cortical thickness, resting state functional connectivity (RSFC)], developmental milestone delays, internalizing symptoms, adverse childhood events).

**Results:** The model distinguishing persistent distressing vs. low-level PLEs showed the highest accuracy (test sample accuracy=69.33%; 95% CI:61.29%-76.59%). The most important predictors included internalizing symptoms, adverse childhood events, and cognitive functioning. Models distinguishing persistent vs. transient distressing PLEs generally performed poorly.

**Conclusions:** Model performance metrics indicated that while most important factors overlapped across models (e.g., internalizing symptoms), adverse childhood events were especially important for predicting persistent distressing PLEs. Machine learning analyses proved useful for distinguishing the most clinically relevant group from the least clinically relevant group but showed limited ability to distinguish among clinically relevant groups that differed in PLE persistence.
Psychotic-like experiences (PLEs) include unusual thought content, including unconventional beliefs, and perceptual abnormalities, including auditory and visual distortions, that lie on the lower end of the psychosis spectrum continuum (1,2). Although PLEs generally refer to non-clinical forms of psychosis spectrum symptoms, individuals experiencing PLEs in childhood and adolescence show greater odds of developing diagnosable mental health concerns, including psychosis (odd ratio=3.96), in adulthood (1). It is important to better understand, as early as possible, which individuals are at risk for the development of mental health concerns versus those that are not at risk. Several studies, including our research using the Adolescent Brain Cognitive Development (ABCD) study data (4), have demonstrated that several characteristics may be particularly important when identifying potentially clinically relevant PLEs, including distress and persistence.

A potential factor distinguishing more benign PLEs and those that transition to psychotic disorders include whether they are persistent, which has been conceptualized as the proneness-persistence-impairment model (PPI model) (3,4). Studies that have examined PLE developmental trajectories (7–10) have generally found evidence for both persistent and transient/relatively benign subgroups (2,9). According to the expanded PPI model, most PLEs may be more benign and transient because exposure to additional environmental risks and/or stronger genetic/early environmental diathesis (potentially reflected in increased expression of pathophysiology over development) is necessary for subclinical psychosis to first become distressing and persistent and, second, to deteriorate into a clinical psychotic state (3).

Risk factors associated with early PLEs include those that reflect relatively stable risk factors present from birth (e.g., family history of psychosis (12)), factors emerging early in life (e.g., delayed developmental milestones, early cognitive deficits (12,13)), as well as factors that may emerge or worsen later in development (e.g., structural and functional neural impairments (14–17), internalizing symptoms, and adverse childhood events (18–20)). Although previous research has implicated this range of risk factors in the development of PLEs, it is unclear which
risk factors are the most important markers of clinically significant PLEs. Previous research examining risk factors has utilized traditional approaches such as univariate models examining each risk factor independent from the other risk factors (11,12,21–27) In contrast, machine learning techniques can capture the multivariate patterns in the data and can identify the most important markers capturing the pattern of data.

The present study represents a follow-up to our previous work using univariate analyses to identify predictors of distress and persistence associated with PLEs(4). The present work expands upon our prior work by aiming to use machine learning models to capture the multivariate patterns of risk factors for early PLEs. Specifically, this study aimed to use random forest classifications to examine the most important predictors of distress and persistence of PLEs measured over three time points when examining a range of multi-modal predictors measured at baseline, with predictors spanning caregiver- and self-report symptoms and experiences, neural metrics, and neurocognition. Random forest classification was chosen as it is robust against overfitting and can include multiple measures simultaneously. Three main models were examined: 1) persistent distressing PLEs versus low-level PLEs; 2) transient distressing PLEs versus low-level PLEs; 3) persistent distressing PLEs versus transient distressing PLEs. Alternative definitions of PLE groups were also examined to test whether findings were specific to our a priori definitions of persistence and distress. These random forest classifications incorporated best practices, including testing models in hold-out samples and tuning parameters(19,20). It is hypothesized that predictors from self/parent-report variables (internalizing symptoms, family history of psychosis, motor and speech developmental delays, adverse childhood experiences) will show stronger associations with distress and persistence of PLEs compared to predictors such as behavior- (cognition), and circuit- (RSFC, structural MRI) levels of analysis that may emerge or worsen later in development. Importantly, the present work identified the most important risk factors at baseline prior to distressing PLEs persisting over time.
Methods

Participants

The ABCD Study is a large-scale study tracking 9-10-years-olds recruited from 21 research sites across the United States (30,31). Potential participants were excluded from participating in the Adolescent Brain Cognitive Development Study for the following reasons: child not fluent in English, MRI contraindication (e.g., irremovable ferromagnetic implants or dental appliances, claustrophobia, pregnant), major neurological disorder, gestational age less than 28 weeks or birthweight less than 1,200 grams, history of traumatic brain injury, or had a current diagnosis of schizophrenia, autism spectrum disorder (moderate, severe), mental retardation/intellectual disability, or alcohol/substance use disorder.

ABCD Data Release 4.0 (DOI 10.15154/1523041) includes 3 full waves of data: baseline (N=11,878), 1-year follow-up (N=11,235), and 2-year follow-up (N=10,416). See Table 1 for sample characteristics. All available data was utilized for measured risk factors (detailed below and in Figure 1), which were obtained at baseline. All procedures were approved by a central Institutional Review Board at the University of California, San Diego. All parents and children provided written informed consent and assent, respectively.

Measures

All measures are described in detail within the Supplemental Methods. For random forest classification analyses, the task was accuracy in distinguishing among: low-level PLEs versus persistent distressing PLEs, low-level PLEs versus transient distressing PLEs, and transient distressing PLEs versus persistent distressing PLEs (see Supplemental Tables 1-2; Supplemental Figure 1 for models examining additional PLE groups). Two types of classification models were trained. The first type, termed “a priori” models, utilized predictors at baseline that were identified in previous work (4,12,32–34): internalizing symptoms, adverse life experiences,
within-network cingulo-opercular and default mode RSFC, average prefrontal cortical thickness, fluid and crystalized cognition composites, parent-reported motor delays, speech delays, and family history of psychosis. The second type, termed exploratory models, included the aforementioned predictors minus the cognitive composites and average prefrontal thickness, as well as all other circuit-level metrics (i.e., thickness and RSFC metrics) for a total of 267 exploratory metrics (7 a priori predictors+218 additional RSFC metrics+35 additional thickness metrics+7 individual cognitive tests; see Supplemental Results).

**Self/Parent Report-Level Measures**

**Prodromal Questionnaire-Brief Child Version (PQ-BC).** Participants completed the previously validated Prodromal Questionnaire-Brief Child Version (PQ-BC)(10). Consistent with previous research(4,12), distress scores were calculated as the total number of 21 questions endorsed weighted by level of distress [i.e., 0=no, 1=yes (but no distress), 2-6=yes (1+score on distress scale); range: 0-126]. See Figure 1 for definitions of PLE groups (see Supplemental Table 1 for alternate definitions of PLE groups).

**Other Psychopathology and Functioning Measures.** Sums of Kiddie-Structured Assessment for Affective Disorders and Schizophrenia (KSADS) for DSM-5(35) youth-rated internalizing symptoms (i.e., number of symptoms of current depression and generalized anxiety disorder) were examined. A history of psychotic disorders was scored as present if the participant had any first- or second-degree relatives with a psychotic disorder history using the Family History Assessment Module Screener(24).

**Developmental Milestones.** The current study examined sums of the total number of caregiver-reported a) motor and b) speech developmental milestone delays(37).

**Adverse Childhood Events.** Based on previous research finding associations with PLEs in the ABCD sample(32), an adverse childhood event (ACE) variable was defined using the number of parent-rated child experience of traumatic experiences from the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS), a parent-rated question from the K-SADS
about whether the child was bullied at school or in the neighborhood, and seven parent-rated questions of financial adversity from a demographic questionnaire (e.g., “Were evicted from your home for not paying the rent or mortgage?”; see Supplemental Methods for all financial adversity questions and methods for creating the ACEs score).

**Behavior-Level Measures**

**Neuropsychological Test Battery.** The current study utilized uncorrected National Institutes of Health Toolbox Cognitive Battery scores from the fluid and crystallized composite scores (38,39).

**Circuit-Level Measures**

**Structural MRI Measures.** All participants completed T1-weighted and T2-weighted structural scans (1mm isotropic) on a 3T scanner (Siemens, General Electric, or Phillips) with a 32-channel head coil. Structural neuroimaging processing was completed using FreeSurfer version 5.3.0 through standardized processing pipelines (40). For the current study, we included average prefrontal cortical thickness, which was the average of the bilateral caudal middle frontal, rostral middle frontal, and superior frontal gyrus (see Supplemental Methods for quality control criteria, n=42 excluded).

**Resting State Functional Connectivity (RSFC).** Participants completed four 5-minute resting-state BOLD scans, with their eyes open and fixated on a crosshair. Resting state images were acquired in the axial plane using an EPI sequence. Other resting-state image parameters varied by 3T scanner and have been previously detailed (https://abcdstudy.org/images/Protocol_Imaging_Sequences.pdf) (33). The data analysis pipeline has also been detailed previously (34,40). Pair-wise correlations were examined for regions of interest (ROIs) within functionally-defined parcellations (i.e., Gordon networks) and subcortical ROIs (40, 41). The Fisher Z-transform of the correlation values was examined for within-network cingular-opercular and within-network default mode ROIs (see Supplemental Methods for quality control criteria, n=87 excluded).
Statistical Analysis

Analyses are pre-registered at:
https://osf.io/y38u4/?view_only=a00d52d8764144b7b476070ef33c0872.

Group Creation. See Figure 1 for group definitions. PLEs in middle childhood may encompass some developmentally appropriate and/or transient experiences and therefore focusing on more clinically relevant features of distress and persistence can enhance the clinical significance of this construct. In accordance with our previously generated definitions of distress and persistence(2), the dichotomous distress factor was coded as one for individuals with >= 1.96 SDs above the mean of distress scores on the PQ-BC, with all other scores coded as zero. Distress thresholds were re-calculated at each wave, to partially account for re-testing effects. These thresholds were chosen based on research using this threshold on different psychosis risk questionnaires in college students(42). A dichotomous persistence factor was coded as one if individuals were coded as a one if individuals met the above definition of distressing PLEs for two or more of the three waves of data (n=305), with transient distressing PLEs coded as zero if they met the definition of distressing PLEs for one time point and were at <= 0.50 SDs above the mean for distressing PLEs for the other two time points (n=374). Participants met criteria for low-level PLEs if they scored <= 0.50 SDs on distressing PLEs on all three waves of data. Originally n=6589 met criteria for the low-level PLEs group, so for analyses comparing low-level PLEs group to other PLEs group(s), the low-level PLEs group was separately matched to the each of the PLE groups in terms of age, sex, race/ethnicity. See Supplemental Table 1 for alternative group definitions, and Supplemental Results for findings using data-driven definitions of PLE groups, created using growth mixture modeling (package lcmm(43)), which were then used to as outcomes in random forest classification models.

Random Forest Classifications. Machine learning techniques, including random forest regression, are becoming more widespread in psychiatric research(44). These types of models have been used in a variety of research questions, including to predict remission of mental
health diagnoses (45), predict future mental health-relevant behavior, such as suicide attempts (46), predict cognitive functioning in psychiatric data (47), as well as to classify psychiatric groups (for a review of how psychiatric research has utilized machine learning techniques, see (44)). Random forest regressions are well suited for psychiatric research in large scale datasets, as this technique can handle the inclusion of many predictors simultaneously and is robust in its predictive accuracy even for many predictors with small effects (48,49). However, random forest regressions are limited in ability to handle within-person correlations, limiting the applicability to longitudinal data.

All random forest classifications were run using the randomforest package in R (50). Missing data was imputed using the rfImpute function (Supplemental Table 3). All predictor variables were scaled using control normalization (51), whereby the predictor variables for the PLE groups were scaled using the average and standard deviations calculated from the predictor variables for the low-level PLEs groups. First, all models were trained on an 80% training sample. The randomforest package uses bootstrapping and averaging to deal with overfitting. An inner nested 10-fold cross-validation loop on the training sample repeated three times was used for tuning parameters (i.e., to find the optimal number of trees and the number of input variables that are randomly chosen at each split; see Supplemental Table 4 for more information about these tuning parameters). Model performance was evaluated in a 20% holdout sample using model accuracy (both overall and for each group). Groups were stratified across training and test samples to ensure equal proportions of groups in every sample. Feature importance was quantified using the randomforest package ‘importance’ function output, the mean decrease in Gini index (i.e., mean decrease in impurity), mean decrease in accuracy index, as well as an additional standard feature importance metric, SHAP (Shapley Additive exPlanations) values, which are considered standard feature importance metrics (52–54). Receiver operating characteristic curves (ROC) were also examined for all models to evaluate the performance of the model at all classification thresholds. The performance of the classifier
was quantified by calculating sensitivity, specificity, negative prediction error and positive prediction error (Table 2). See Supplemental Results for follow-up analyses examining distress and persistence as factors; see Supplemental Tables 2 and Supplemental Figure 1 for results with alternative PLE definitions.

Results

*A priori* models performed similarly to exploratory models (see Supplemental Results). Accordingly, we will focus on the results obtained by the *a priori* models.

**Persistent Distressing PLEs Distinguishable from Low-level PLEs But Not Transient Distressing PLEs**

*Distinguishing Persistent Distressing vs. Low-level PLEs*. The model predicting persistent distressing versus low-level PLEs showed fair fit, including an out of box error rate of 34.43% (31.69% for low-level PLEs, 37.19% for persistent distressing PLEs; see Figure 1 for a standard ROC and Area Under the Curve (AUC) for the model; see Table 2 for model characteristics). The model in the test sample showed an accuracy of 68.03% (95% CI: 58.98%-76.18%). Several predictors were consistently influential in this model across feature importance metrics (i.e., mean decrease in accuracy, mean decrease in Gini, SHAP values) including ACEs, internalizing symptoms, and fluid and crystalized cognition composites (see Figure 1).

*Distinguishing Transient Distressing vs. Low-level PLEs*. As can be seen in Figure 1, the model predicting transient distressing versus low-level PLEs performed poorly, showing a worse fit than the persistent distressing versus low-level PLEs model, including an out of box error rate of 43.98%, which was greater for the transient distressing PLE group (30.10% for low-level PLEs, 57.86% for transient distressing; see Table 2 for model characteristics). The model in the test sample showed an accuracy rate of 61.33% (95% CI: 50.05%-69.16%). There was not consistency across feature importance metrics, mean decrease in Gini and SHAP values showed within-network cingulo-opercular and default mode network RSFC as particularly
influential features, whereas mean decrease in accuracy scores showed that internalizing symptoms was particularly influential.

**Distinguishing Persistent Distressing vs. Transient Distressing.** The model predicting persistent distressing versus transient distressing showed a relatively poor fit, including an out-of-box error rate of 47.32%, which showed some imbalance across groups (27.27% for persistent distressing, 63.55% for transient distressing PLEs) (see Figure 1; see Table 2 for model characteristics). The model in the test sample showed an accuracy of 63.97% (95% CI: 55.30%-72.02%). Several predictors were consistently influential in this model across feature importance metrics (i.e., mean decrease in accuracy, mean decrease in Gini, SHAP values) including ACEs, internalizing symptoms (see Figure 1).

**Models using Alternative PLE Group Definitions.** The Supplemental Results contain findings using a number of alternative definitions of PLE groups (i.e., outside of the original persistence and transient distressing PLE group definitions). With these alternative definitions, the highest accuracy was found if constraining the persistent distressing group to those that met criteria at all three waves (or at least the first and third wave; see Supplemental Table 2). Importantly, regardless of PLE group definition, the most important factors distinguishing higher PLE groups from lower PLE groups were internalizing symptoms and ACEs, with some evidence for cognitive factors showing greater importance relative to other factors (see Supplemental Figure 1).

**Discussion**

The present study examined whether PLEs examined over three assessment waves could be classified based on baseline metrics. Results indicate that persistent distressing PLEs can be distinguished from low-level PLEs, but not transient distressing PLEs. The models appeared to have a gradient in accuracy, whereby persistent distressing versus low-level PLEs showed the greatest accuracy, followed by transient distressing versus low-level PLEs, with the persistent distressing versus transient distressing showing slightly lower accuracy. Models
generally indicated the importance of baseline internalizing symptoms, adverse childhood events, and cognitive metrics in classifying worsening PLEs. Baseline adverse childhood events was particularly important for distinguishing persistent distressing PLEs. Although several metrics (baseline distressing PLEs, internalizing symptoms, adverse childhood events, and cognitive function) may be particularly important for distinguishing worsening PLEs, overall baseline metrics struggled to classify PLE groups, such as persistent from transient PLEs. Importantly results remained consistent regardless of PLE group definition (Supplemental Results). More nuanced predictors, perhaps incorporating changes over time, are likely needed to effectively classify persistent from transient PLEs. PLEs are a heterogenous construct, with only a subset at risk for the development of a psychosis spectrum illness, although others developing other mental health conditions (55). Given this heterogeneity of overall endorsement of PLEs, the results point to the importance of including factors that increase the clinical relevance of these endorsements, including persistence over time, when examining PLEs in large general population samples.

**Classifying Persistent Distressing PLEs**

The model predicting persistent distressing from low-level PLEs generally performed fair/average. The persistent distressing versus low-level PLEs models showed relatively balanced sensitivity and specificity in the test sample, indicating the presence of both false positive and false negatives. Likewise, there was generally balanced categorization errors across persistent distressing and low-level PLEs groups. For differentiating the persistent distressing PLEs group from low-level PLEs group, it appears that several markers consistently predicted persistent distressing PLEs across *a priori* or exploratory models (see Supplemental Results for exploratory models). These factors included greater internalizing symptoms, greater ACEs, worse crystalized and fluid cognition, with exploratory models finding evidence for specifically picture vocabulary and list sorting working memory tests. This is consistent with our previous work on persistent distressing PLEs (2), and points to these factors as particularly
useful predictive markers. Finding that internalizing symptoms was an important marker may be in part due to overlap with distress associated with PLEs (44), perhaps pointing to the role of distress and/or negative affect in classifying clinically significant PLEs. These results remained consistent regardless of how PLE group was defined, including if examining baseline predicting 2-year follow-up PLEs, and when focusing one or more significantly distressing PLEs (a PLE rated in the 3-5 range on the distress scale) endorsed for two or more of the last three assessment waves (Supplemental Results). Overall, persistent distressing PLEs were most distinguishable from low-level PLEs, highlighting both the severity of the persistent distressing PLE group, but also the difficulty in distinguishing amongst PLE groups.

**Classifying Transient Distressing PLEs**

The transient distressing PLEs versus low-level PLEs models performed poorly and performed worse than the persistent distressing PLEs versus low-level PLEs models. In models classifying transient distressing PLEs, neural metrics (e.g., CON and DMN within-network RSFC) showed evidence of being important metrics (according to SHAP values and mean decrease in Gini). This is consistent with our previous research indicating that transient distressing PLEs showed stronger neural impairments compared to low-level PLEs(4), and may point to distressing PLEs, regardless of persistence, showing pathophysiological impairments. Difficulty distinguishing transient distressing PLEs from other groups may reinforce the notion that transient groups may be particularly heterogeneous in terms of clinical presentation, and therefore may have shown greater variability in terms of most important predictors (41,42).

*Can we predict persistent from transient distressing PLEs?*

Models classifying persistent distressing PLEs from transient distressing PLEs based on baseline metrics generally performed poorly. This is perhaps not entirely surprising, given the
small effect-sized differences distinguishing these groups in previous research using the ABCD Study (2,52) and other samples (51). Although previous work has used machine learning to distinguish between psychotic disorder groups on neural metrics with accuracy upwards of 75% (53), consistent with the present work there is limited evidence for the ability to use these methods to predict features from subclinical PLE scores (54) Along these lines, analyses from the present work showed few metrics consistently distinguishing persistent distressing from other PLE groups, but these metrics included ACEs, highlighting the importance of stressful life events. Evidence that ACEs was especially important for persistent distressing PLEs is consistent with the persistence proneness impairment (PPI) model, which indicates that it is only in the context of greater environmental exposures that risk translates into worsening psychosis spectrum symptoms (55). Previous research indicates that traumatic life events may be an important predictor of worsening psychotic symptoms, including up to 80% of individuals at high risk for psychosis endorsing a traumatic event in childhood (56).

Overall, the performance indices, including accuracy, from models classifying persistent from transient distressing PLEs indicate there is little evidence that these PLE groups can be distinguished from one another based on the examined baseline metrics. It may be the case that a) other factors beyond baseline metrics, such as changes in functioning or symptoms, b) currently unmeasured phenotypes, or c) there may be no phenotypes that can accurately distinguish between PLEs groups. Future research should incorporate changes in risk factors. Additionally, to the second point about the role of unmeasured phenotypes, exploratory models failed to improve prediction and often found that the variables included in the models presented in this manuscript (e.g., internalizing symptoms, ACEs, fluid cognition) were still amongst the most important predictors. To the last point, it is possible that no cross-sectional self-report, behavioral, or circuit-level predictors may accurately classify PLEs groups. Notably, circuit-level predictors were not generally amongst the most important predictors for distinguishing persistent distressing PLEs. This may indicate that either these predictors don’t distinguish
between these groups after accounting for behavioral-level predictors, or it may point to the need for more precise neural predictors. Further, the overall results may point to the heterogeneity of early PLEs, with this heterogeneity perhaps masking the ability to uncover unifying risk factors. The heterogeneity of transient PLEs is highlighted by the poor accuracy of these groups regardless of how defined (see Supplement). Therefore, it may not be until the worsening of symptoms or onset of functional impairment that symptoms are of sufficient severity to be distinguished from less severe PLEs.

Limitations

The present work has important contextualizing factors, including that the models generally showed fair to poor accuracy and therefore are likely not currently useful in classifying amongst clinically relevant PLE groups, only classifying persistent distressing PLEs from low-level PLEs. It is possible that as future waves of ABCD data are released and group classifications can incorporate additional waves of data, persistent and transient PLE groups classifications will show greater accuracy. Also, the analyses examined risk factors at baseline prior to the development of persistence of PLEs, and therefore there is overlap in the time points examined for risk factors and PLEs. Prior to using these markers for screening purposes, efforts should be made to minimize false negatives (i.e., maximize sensitivity) (57) as the risks associated with missing anyone experiencing persistent distressing PLEs outweigh the risks of identifying a false positive. Machine learning models involve a number of choices, including decisions to use hold out samples, size of training and test sets, and choice of type of machine learning model. The present work aimed to use best practices to aid in determining the most important predictors of persistent distressing PLEs (19,20).

Conclusions

Analyses indicate that machine learning analyses can be used to distinguish persistent distressing from low-level PLEs groups. Additional evidence indicated that distress may be more
distinguishable than persistence, including evidence that models performed poorly when trying to distinguish persistent from transient distressing PLEs. Regardless, the work also reinforces the importance of several metrics that have been previously implicated in the development and worsening of early psychotic experiences. Adverse childhood events appeared especially important for persistent distressing PLEs, potentially providing support for the PPI model – that worsening PLEs is linked to greater environmental loading. Both persistent and transient distressing PLEs were linked to risk factors including internalizing symptoms. Important, these results were robust to PLE group definition. Overall, evidence indicates that while persistent distressing PLEs are distinguishable from low-level PLEs, there is limited evidence for the ability to distinguish between clinically relevant PLEs groups.
Acknowledgments

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcodstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 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 U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093, and U01DA041025. A full list of supporters is available at https://abcodstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcodstudy.org/Consortium_Members.pdf. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in 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 DOI 10.15154/1519007.

Funding/Support: This work was supported by National Institute on Drug Abuse (U01 DA041120 to D.M.B.); National Institute of Mental Health (K23MH121792-01 to N.R.K.); BrightFocus Foundation (A.S.).

Additional Acknowledgments: We thank the families participating in the Adolescent Brain and Cognitive Development study.
Financial Disclosures

Dr. Barch, Dr. Jackson, and Dr. Karcher receive research funding from the National Institute of Mental Health. Dr. Barch receives research funding from the National Institute on Drug Abuse. Dr. Sotiras receives research funding from the BrightFocus Foundation. Dr. Jackson receives research funding from the National Institute of Neurological Disorders and Stroke. Dr. Niendam is a co-founder and shareholder in Safari Health, Inc. Dr. Walker reports no biomedical financial interests or potential conflicts of interest.
References


Table 1. Participant Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Persistent Distressing vs. Low-level PLEs Models</th>
<th>Transient Distressing vs. Low-level PLEs Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Persistent Distressing (n=305)</td>
<td>Matched Low-level PLEs (n=305)</td>
</tr>
<tr>
<td>Total ABCD Sample (n=11,878)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Sex (% female)</td>
<td>47.8</td>
<td>5683</td>
</tr>
<tr>
<td>Race/Ethnicity, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>52.1</td>
<td>6182</td>
</tr>
<tr>
<td>Black</td>
<td>15.0</td>
<td>1784</td>
</tr>
<tr>
<td>Asian</td>
<td>252</td>
<td>10.5</td>
</tr>
<tr>
<td>Multiracial/ Multiethnic</td>
<td>1247</td>
<td>10.5</td>
</tr>
<tr>
<td>Hispanic</td>
<td>20.3</td>
<td>2411</td>
</tr>
<tr>
<td>Age at Baseline</td>
<td>9.940</td>
<td>0.624</td>
</tr>
<tr>
<td>Baseline Distressing PLEs</td>
<td>6.317</td>
<td>10.607</td>
</tr>
<tr>
<td>1-Year Distressing PLEs</td>
<td>4.617</td>
<td>9.380</td>
</tr>
<tr>
<td>2-Year Distressing PLEs</td>
<td>3.562</td>
<td>7.721</td>
</tr>
</tbody>
</table>

Abbreviations. PLEs=Psychotic-like experiences

aTotal ABCD Sample included for comparison purposes. Persistent Distressing PLEs group showed PLE distress scores $\geq 1.96$ standard deviations (SDs) above the mean for 2+ assessment waves; Transient Distressing PLEs showed PLE distress scores $\geq 1.96$ standard deviations (SDs) above the mean for 1 assessment wave and PLE distress scores $\leq 0.50$ SDs below the mean for the other 2 waves; Low-level PLEs group showed PLE distress scores $\leq 0.50$ SDs below the mean for all 3 assessment waves. Matched low-level PLE groups were matched to the PLE groups based on age, sex, and race/ethnicity.
Table 2. Model Characteristics

<table>
<thead>
<tr>
<th>PLE Groups</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive Predictive Value</th>
<th>Negative Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistent Distress v. Low-level PLEs</td>
<td>0.623</td>
<td>0.738</td>
<td>0.704</td>
<td>0.662</td>
</tr>
<tr>
<td>Transient Distressing v. Low-level PLEs</td>
<td>0.707</td>
<td>0.520</td>
<td>0.596</td>
<td>0.639</td>
</tr>
<tr>
<td>Persistent Distressing v. Transient Distressing</td>
<td>0.295</td>
<td>0.920</td>
<td>0.75</td>
<td>0.616</td>
</tr>
</tbody>
</table>

Abbreviations. PLEs=psychotic-like experiences.
Figure Legends

Figure 1. Summary of findings from random forest classifications performing three tasks distinguishing: persistent distressing psychotic-like experiences (PLEs) from low-level PLEs, persistent distressing from transient distressing PLEs, and transient distressing from low-level PLEs. (A) Mean decrease in accuracy for each of the predictors for each of the tasks, (B) Mean Decrease in Gini coefficient for each of the predictors for each of the tasks, (C) Mean SHAP values for each of the predictors for each of the tasks, and (D) Receiver Operating Characteristic Curves (ROC) and Area Under the Curve (AUC) for each of the tasks. Note, for (A), (B), and (C), the dimensions of size and color both indicate the relative importance of the input feature, such that features with larger red circles reflect greater importance.
Mean Decrease in Gini for Each of the Tasks

(A)

PD vs. LP
PD vs. TD
TD vs. LP

Internalizing Symptoms
ACEs
Fluid Cognition
Crystalized Cognition
Prefrontal Thickness
CON RSFC
DT RSFC

Mean Decrease in Accuracy for Each of the Tasks

(B)

PD vs. LP
PD vs. TD
TD vs. LP

Internalizing Symptoms
ACEs
Fluid Cognition
Crystalized Cognition
Prefrontal Thickness
CON RSFC
DT RSFC

Mean SHAP Values for Each of the Tasks

(C)

PD vs. LP
PD vs. TD
TD vs. LP

Internalizing Symptoms
ACEs
Fluid Cognition
Crystalized Cognition
Prefrontal Thickness
CON RSFC
DT RSFC

Group Definitions

PD = Persistent Distressing (high threshold [≥1.96SD above the mean of distressing PLEs] for ≥ 2 assessment waves)

TD = Transient Distressing (high threshold for 1 assessment wave, low threshold [≤0.5SD above the mean of distressing PLEs] for 2 assessment waves)

LP = Low-level PLEs (low threshold for all 3 assessment waves)