Genome-wide Polygenic Scores for Multiple Psychiatric and Common Traits Identify Preadolescent Youth with Risk for Suicide

Yoonjung Yoonie Joo

Department of Psychology, Seoul National University

Seo-Yoon Moon

College of Liberal Studies, Seoul National University

Hee-Hwan Wang

Department of Psychology, Seoul National University

Hyeonjin Kim

Department of Psychology, Seoul National University

Eun-Ji Lee

Department of Psychology, Seoul National University

Jonathan Posner

Department of Psychiatry, Columbia University Medical Center

Woo-Young Ahn

Department of Psychology, Seoul National University

Incheol Choi

Department of Psychology, Seoul National University

Jae-Won Kim

Department of Child and Adolescent Psychiatry, Seoul National University Hospital

Jiook Cha (✉ connectome@snu.ac.kr)

Department of Psychology, Seoul National University <https://orcid.org/0000-0002-5314-7992>

Research Article

Keywords: Psychiatric Genetics, Genome-wide Polygenic Score, Childhood Suicidality, Early Life Stress

Posted Date: September 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-136949/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Importance.

Suicide is the second leading cause of death in children worldwide but no available means exist to identify the risk in youth.

Objective.

To predict the risk of suicide in children and to investigate whether and to what extents genetic factors and a major environmental risk factor, early life stress(ELS), influence youth suicide.

Design, Setting and Participants.

We analyzed the genotype-phenotype data of 11,869 preadolescent children ages 9- to 10-year-old from the Adolescent Brain and Cognitive Development (ABCD) study. We estimated genome-wide polygenic scores (GPSs) of 25 complex traits to investigate their phenome-wide associations and predictive utility with suicidality (suicidal ideation and attempt) with machine learning approaches.

Predictors.

GPSs of 25 traits including psychiatric disorders, personality, cognitive capacity, and psychological traits. Parent Child Behavior Checklist to measure ELS in youth and Youth Family Environment Scale to assess family environment.

Main outcomes and Measures.

Records of suicidal ideation and attempt of the participants were derived from the computerized version of Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS).

Results.

We identified three GPSs associated with youth suicidality in multiethnic (n = 7,206) and European-ancestry (n = 5,749) participants: ADHD ($P = 3.48 \times 10^{-4}$; odds ratio = 1.13 in multiethnic participants, $P = 5.60 \times 10^{-5}$, OR = 1.25 in European-ancestry participants), general happiness ($P = 1.43 \times 10^{-3}$; OR = 0.89 in multiethnic, $P = 8.61 \times 10^{-4}$, OR = 0.89 in European) and autism spectrum disorder(ASD) ($P = 1.81 \times 10^{-3}$; OR = 1.15 in multiethnic, $P = 1.26 \times 10^{-3}$, OR = 1.18 in European). We also found a significant GPS-by-environment interaction between the effects of genetic risk factors for ASD and the level of ELS in

increasing the risk for suicidal ideation ($P = 1.36 \times 10^{-2}$, OR = 1.12 in multiethnic, $P = 1.39 \times 10^{-3}$, OR = 1.19 in European). A machine learning model trained on the same data showed moderately accurate prediction of children with overall suicidal ideation with a test ROC-AUC of 0.727 (0.746 in European), and with suicidal attempts with a test ROC-AUC of 0.641 (0.975 in European) in held-out samples.

Conclusions and Relevance.

This study provides the first quantitative account of polygenic and environmental factors of suicidality in a large, representative population of preadolescent youth. It thus shows the potential utility of the GPSs in identifying a child with high risk for suicidality for early screening, intervention, and prevention.

Introduction

Every 1 minute 2.6 persons attempts suicides and succeeds every 11 minutes in the U.S¹ In youth, suicide has been the 2nd leading causes of death worldwide; the suicide-rate curve has not been reduced in decades.^{1,2} Identifying a child with risk for suicide is challenging and existing approaches show poor predictive validity.³

Suicidal behaviors, including suicidal ideation and/or attempt, are under significant genetic influences as demonstrated in twin and family studies^{4,5}, with the estimated heritability ranging from 30% to 55%^{4,6}. The literature supports that an inherited genetic factor may account for suicidal risk, and furthermore, the genetic predisposition may be aggravated by psychiatric illnesses and environmental factors such as stress⁷. No literature has yet to report this in children.

The genetic architecture of suicidal behaviors is known to be polygenic^{8,9}, being influenced by the cumulative effects of numerous single-nucleotide polymorphisms (SNPs) each with miniscule effects. Recent Genome-wide Association Studies (GWAS) revealed several loci linked to completed suicide^{10,11}, suicide attempt^{12,13}, and suicidal ideation^{14,15,16,17}. But none of the loci have been replicated across studies. Based on the findings from GWAS, genome-wide polygenic scores (GPS) integrate cumulative effects of genome-wide SNPs^{18,19} and has emerged as a promising tool to efficiently predict the likelihood of a complex trait from a genomic perspective.^{20,21} GPSs can be used to stratify individuals at higher biological risk or to provide insights into the possible genetic overlaps among complex traits, potentially enabling the goal of personalized medicine.^{22,23,24}

Genetic predisposition of suicidal behaviors has been reported to interact with environmental factors (GxE interaction)⁹, such as early life stress (ELS),^{25,26} potentially via epigenetic mechanisms^{27,28}. Investigating whether and to what extents ELS and genetic factors together act synergistically on youth suicide will offer a much-needed insight into the biological pathway of suicide and an actionable target for intervention. To date, no studies have reported this GxE interaction or the predictive utility of genetic

factors for youth suicidality. In this study, we investigate the extent to which GPSs^{22,23} for common traits and psychiatric disorders, and their interaction with ELS is linked to the risk for suicide in young children.

Methods And Materials

Study design and participants

The study sample includes 11,869 preadolescent children ages of 9 to 10 years old in the US across 21 sites recruited between 2015 and 2019 from the Adolescent Brain and Cognitive Development (ABCD) cohort. Among the initial samples, 1,658 cases and 124 cases were identified for suicidal ideation and suicide attempt, respectively. Ethical approval for the study was obtained from Seoul National University IRB. Full details of the study, measures, and samples can be found elsewhere.²⁹

Genotype data

Saliva samples of the participants were collected and genotyped at Rutgers University Cell and DNA Repository using Affymetrix SmokeScreen Array consisting of 733,293 single nucleotide polymorphisms (SNPs). After removing the SNP with genotype call rate <95%, sample call rate <95%, and minor allele frequency (MAF) <1%, raw genotypes were imputed toward 1000G Phase3 reference panel using the Michigan Imputation Server³⁰ and phased with Eagle v2.4³¹ (a total of 12,046,090 variants). Additional quality control (QC) process removed SNPs with INFO score <0.4, genotype call rate <95%, Hardy-Weinberg Equilibrium p-value <1E-20 for ethnically diverse population, sample missingness >5%, and minor allele frequency (MAF) <0.5%. We also removed samples with extreme heterozygosity having F coefficient bigger than 3 standard deviations of the population mean.

Our study samples have complex genetic structures owing to the presence of related family members, diverse genetic ancestries, and admixed samples innate to the U.S. population. We used PC-Air³² to estimate ancestrally-informative PCs of the genotypes that is robust to the related pedigree structure, and PC-Relate³³ to provide accurate estimate of recent genetic relatedness measures from the population with ancestry admixture. The details of QC procedure are available in **Supplementary Method**. In the rest of this paper, we used the genotype data (a total of 11,301,999 variants) of 8,760 unrelated samples after QC, and the final PCs were used as covariates in all the reported analyses.

Construction of Genome-Wide Polygenic Scores (GPS).

For GPS generation, we selected 25 psychiatric and common traits that are known to be related to suicidality and have their GWAS summary statistics data publicly available, including personality, cognitive, psychological traits, and psychiatric disorders that are known to be broadly related to

suicidality: general happiness,^{34,35} insomnia,³⁶ depression,³⁷ risk behaviors,³⁸ risk tolerance, educational attainment,^{39,40} cognitive performance,^{38,39} snoring,⁴¹ worry,⁴² IQ,³⁹ cannabis usage,^{43,44} drink per week,⁴⁵ smoker,⁴⁶ Attention deficit hyperactivity disorder (ADHD),⁴⁷ Autism spectrum disorder (ASD),⁴⁸ major depressive disorder (MDD),⁴⁹ schizophrenia,⁵⁰ bipolar disorder,⁵¹ post-traumatic stress disorder (PTSD)⁵² and Alzheimer's disease⁵³ For general happiness, four GPS were built and tested for the study: the participants were asked with different questionnaires about their level of subjective well-being, such as (1) two different GWAS of “how happy are you in general”, (2) “how happy are you with your health in general”, and (3) “to what extent do you feel your life to be meaningful”. All the GWAS summary statistics of the aforementioned traits are publicly available and have been collected for GPS generation.

We performed clumping and pruning of SNPs using PRSice2⁵⁴ with a clumping window of 500 kb, clumping r^2 of 0.2 based on the 1000 Genomes cosmopolitan panel, and no thresholding of p-value significance on the summary statistics since we wanted to fully incorporate the effects of all the SNPs. The GPS for each individual was then computed as a sum of their SNPs adjusting for first 10 genotype PCs, with each SNP being weighted by the effect in the discovery samples.⁵⁴ Since most of the summary statistics were derived from the European population, we selected 5,749 European-ancestry participants (genetic ancestry determined with fastSTRUCTURE algorithm⁵⁵, available from ABCD release 3.0) and performed the same analysis to determine the transferability of Eurocentric PRS in different ethnic groups.

Outcomes and Measures

Suicidal ideation and attempt data were derived from the computerized version of Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS).^{56,57} Passive suicide ideation is wanting to be dead and active ideation is considering suicide with specific methods or plans. Among parent and child reports of K-SADS, we used the one that reported more severe symptoms or diagnoses. The following variables were additionally considered for inputs to the classification models. For psychopathology, ABCD Parent Child Behavior Checklist (CBCL); for intelligence, NIH toolbox;⁵⁸ for family environment, Youth Family Environment Scale; and for ELS, abuse (physical and sexual), household challenges (family substance abuse, family mental illness, family criminals, parental separation or divorce, and violently treated mothers) and neglect (emotional and physical neglect). The items assessing ELS were extracted from various measurement tools reported by children themselves or parents (**Table S1**). An ELS composite score was calculated by averaging z-standardized scores of each subtype. Higher composite score indicates more severe ELS.

Statistical Analysis and Machine Learning Prediction

After QC and GPS construction, the complete dataset of phenotypic outcomes, GPS, and covariate data were available for 7,206 children and used for statistical analysis (**Figure 1**). Associations of the GPSs with suicidal variables within the ABCD children were tested using logistic regression with the following covariates: sex, age, site for sample collection, and self-reported race. We use false discovery rate (FDR) to control for multiple statistical comparisons. We also performed the same analysis after excluding any control individuals with one or more K-SAD records (n=3,108 healthy controls remained).

For suicidal risk prediction with GPSs, we trained CatBoost⁵⁷ model, a recent high-performance gradient boosting model on decision trees. The following input features were used: 25 GPSs, socio-demographic information (sex, age, self-reported ethnicity, study site), psychological information (CBCL), family environment factors (Youth Family Environment Scale), and ELS. We rationalized that the multiple GPSs would account for the multi-dimensional genetic predisposition to suicidal behaviors. For model training and evaluation, we split the data into 80% and 20% in which samples were balanced (bootstrapping without replacement). Within the 80% training data, we conducted five-fold stratified cross-validation and optimized hyperparameters in cross-validation sets. Since the training data have an imbalanced count of case and control, we undersampled every cross-validation set to make an equal number of cases and controls.

Results

In this study, we leveraged a nationwide prospective cohort study of 11,869 preadolescent children ages of 9 to 10 years old in the US across 21 sites recruited between 2015 and 2019. After QC of genotype and phenotype data, the study analyzed 7,206 multiethnic samples including a total of 1,020 identified cases with suicidal thoughts or behaviors: 1,014 cases with suicidal ideation and 71 cases with suicidal attempts were identified with the K-SADS (**Figure 1**). Preadolescent children with suicidal ideation or attempts showed similar sociodemographic, behavioral, clinical characteristics to controls without suicidal ideation or attempts (**Table 1**).

We found several GPS significantly associated with suicidality. Out of the 25 GPSs, a greater GPS of ADHD significantly correlated with a greater likelihood of suicidal ideation ($P=3.48 \times 10^{-4}$, odds ratio (OR)=1.13; FDR significance) from the analysis of the entire 7,206 multiethnic participants. The ADHD GPS significantly correlated with overall suicidality ($P=4.64 \times 10^{-4}$, OR=1.13), active suicidal ideation ($P=4.71 \times 10^{-4}$, OR=1.17) and passive suicidal ideation ($P=1.29 \times 10^{-3}$, OR=1.13) (**Table 2(a)**, **Figure 2**). In active suicidal ideation, a greater GPS of ASD correlated with a greater likelihood of suicidal ideation ($P=1.81 \times 10^{-3}$, OR=1.15). Conversely, a smaller GPS of general happiness significantly correlated with a greater likelihood of passive suicidal ideation ($P=1.43 \times 10^{-3}$, OR = 0.90) and overall suicidal ideation ($P=1.47 \times 10^{-3}$, OR = 0.90). Correlations among the tested GPSs were examined and none of their correlations seemed to affect the results (**Figure S1**).

Analysis of the participants of European descent only (n=5,749) revealed of the GPSs examined the ADHD GPS had the strongest association with all types of suicidality (strongest with overall suicidal

ideation, $P=1.43 \times 10^{-3}$, OR = 0.90). Significant inverse associations between suicidal behaviors and general happiness GPS (strongest with passive suicidal ideation, $P=8.61 \times 10^{-4}$, OR = 0.87) were found with negative effect sizes. The European-only analysis identified additional GPSs correlated with youth suicidality compared to the analysis of the entire multiethnic participants: Smoker GPS (strongest with overall suicidality, $P=1.08 \times 10^{-3}$, OR = 1.13), MDD GPS (strongest with overall suicidality, $P=2.65 \times 10^{-3}$, OR = 1.12), risky behaviors GPS (strongest with passive suicidal ideation, $P=3.43 \times 10^{-3}$, OR = 1.13), and schizophrenia GPS (with suicide attempt, $P=3.55 \times 10^{-3}$, OR = 1.51) (**Table 2(b)**).

Of those three GPSs (ADHD, ASD, general happiness) showing significant associations with suicidality, we further tested their interaction with ELS. We found significant ASD-GPS-by-ELS interaction on active suicidal ideation in both European and multiethnic analyses: a greater ASD GPS in the presence of ELS correlated with a synergistic increase in these likelihood of suicidality ($P=1.36 \times 10^{-2}$, OR=1.12 in entire participants, $P=1.39 \times 10^{-3}$, OR=1.19 in European-ancestry participants, FDR significant). Effects were adjusted for the same covariates with the regression analysis. Bivariate associations between ELS and suicidality were non-significant ($P_s > 0.23$). When stratified by sex, no significant associations were found.

Additionally, as a sensitivity analysis, we limited our analysis to healthy children, defined as participants with no active or past records on the K-SADS ($n=3,108$). The association between suicidality and the follow GPSs remained significant or even became slightly stronger in terms of effect size: ADHD GPS (strongest association with overall suicidal ideation, $P=1.36 \times 10^{-4}$, OR=1.17), ASD GPS (strongest association with active suicidal ideation, $P=5.78 \times 10^{-4}$, OR=1.19) and general happiness GPS (strongest association with overall suicidal ideation, $P=3.21 \times 10^{-3}$, OR=0.89) (**Table S2(a)**). We also found additional GPS associations with youth suicide, including smoker GPS (strongest association with overall suicidal ideation, $P=9.74 \times 10^{-4}$, OR=1.14), PTSD GPS (strongest association with overall suicidality, $P=2.15 \times 10^{-3}$, OR=1.13), MDD GPS (strongest association with overall suicidality, $P=3.77 \times 10^{-3}$, OR=1.13), all of which were identified in the analysis of European-only analysis. These significant GPS associations stayed consistent in the analysis of European-ancestry participants with healthy controls who have no K-SADS record (**Table S2(b)**).

We next tested whether machine learning and the consideration of multiple GPSs could contribute to the prediction of youth suicidality. In predicting suicidal ideation, the 5-fold cross-validated and optimized CatBoost⁵⁷ classifier model showed an ROC-AUC of 0.716, accuracy of 0.601, sensitivity of 0.842, specificity of 0.357, positive predictive value of 0.570, negative predictive value of 0.691 in the balanced held-out test set ($N= 366$; bootstrap samples) (**Figure 3**). Adding the GPS to the phenotype model resulted in a 1.1% boost in ROC-AUC (0.727). In predicting suicidal attempts, the cross validated and optimized stacked ensemble model showed an ROC-AUC of 0.756, accuracy of 0.640, sensitivity of 0.923, specificity of 0.333, positive predictive value of 0.600, negative predictive value of 0.800 in the balanced held-out test set ($N= 25$; bootstrap samples) (**Figure 3**). When predicting based only on participants of European-ancestry, the model performance increased in ideation, but more dramatically in attempt. The phenotype

model showed ROC-AUC of 0.716 in ideation and 0.736 in attempt. In presence of GPS, ROC-AUC increased to 0.746 in ideation and 0.975 in attempt.

Discussion

This study provides the first quantitative account of the polygenic and environmental factors of youth suicidality in a large-scale, representative, multi-ethnic sample of preadolescent children. We found novel associations between youth suicidality and GPS for ADHD, ASD – positive associations – and general happiness – a negative association. We next found significant gene-by-environment interactions between the GPS of ASD and ELS,²⁷ a known risk factor for youth suicidality, together acting cumulatively on youth suicidality. In our predictive modeling, together with the self-reported questionnaires for psychopathology (CBCL), cognitive intelligence (NIH Toolbox), family environment, and socio-demographic variables, inclusion of the multiple GPSs permitted moderately accurate identification of preadolescent youth with suicidality.

We identified significant associations between youth suicidality and the genetic components of psychiatric disorders, i.e., ADHD and ASD, particularly relevant to childhood psychopathology, and of a common trait, i.e., general happiness. For ADHD and ASD, a greater polygenic score correlated with a greater likelihood of suicidal ideation, whereas for general happiness, a smaller polygenic score correlates with a greater likelihood of passive suicide ideation. This is in line with the literature, suggesting that ADHD is associated with suicidal attempts in youth.⁶⁰ Our GPS results highlight possible genetic overlaps between ADHD and youth suicidality. The explained variance (McFadden's pseudo- R^2)⁶¹ of suicide attempt by ADHD GPS was approximately 6.6% in multiethnic children (6.2% in European-ancestry children). This estimation is higher than the results from past PRS studies of suicidality, which showed maximum variance explained by 0.13-0.20% with self-harm behaviors⁶² or up to 0.30-0.70% of the phenotypic variance for suicide attempt explained by depression-based GPS.^{63,64}

The GPS of ASD not only correlated with suicidal ideation, but also showed a significant interaction with ELS. These results corroborate previous findings of preadolescent youth with ASD being at an increased risk for suicidality.^{65,66} One study shows that youth with ASD are 28 times more likely to endorse suicidal thoughts or behaviors than their unaffected peers.⁶⁷ This strong link between risk of suicidality and elevated autistic traits could be explained by behavioral attributes of both phenotypes, such as poor socialization and problem-solving skill, or increased levels of impulsivity and anxiety.⁶⁵ Although literature has reported several social risk factors underlying the ASD-suicidality link,^{66,67} the genetic and environmental contribution to the overlap of ASD and suicidal behaviors remained unknown.⁶⁹ Our analysis shows that the interaction between adverse childhood experience and the genetic risk for ASD is associated significantly with the risk for suicidality. To our knowledge, this is the first report of the gene-by-environment interaction linking ASD to suicidality. These results may imply that for those who has the genetic liability for ASD an adverse childhood experience is particularly risky with regard to suicidality. This might suggest a potential actionable target for intervention.

In line with our findings, previous literature has reported high suicidal risk in individuals who in addition to ASD also have ADHD⁶⁹. As 20-50% of individuals with ASD are known to have comorbid ADHD,⁷⁰ our significant results of ASD and ADHD GPSs associated with suicidality suggest the genetic liability of psychiatric comorbidity for high suicidal risk in children.

Our study suggests that higher genetic liability for general happiness (i.e., the tendency to believe that her or his life is meaningful) underlies a decreased risk of suicidality in youth. Of the several measures of subjective well-being (e.g., happy in general, happy with your health in general, etc), only the GPS for the belief of own meaningful life is significantly linked to youth suicidality. Prior studies report a negative correlation between subjective happiness and suicide^{34,35}. Due to the absence of relevant subjective well-being variables in the ABCD study, we could not further investigate their phenotypic correlation.

Our findings provide evidence for utilizing multiple GPSs for suicide prediction and contributes to our understanding of its genomic architecture. The combination of cross-phenotype GPSs, family environment, behavioral, psychological scales, and ELS allowed accurate prediction of suicidality with a ROC-AUC up to 0.727 (0.746 in European-ancestry only) for suicidal ideation, and 0.641 (0.975 in European-ancestry only) for suicidal attempts on the balanced test set in the entire multiethnic samples. Previous studies tested different methods for suicide risk prediction and showed substantial classification ability having AUC ranging from 0.74 to 0.88, using known sociodemographic and psychiatric risk factors.^{71–74} Although successful, some of the previous models included long-term measurement of risk indices such as 12-month or lifelong risk factors, which are not readily measured in naturalistic clinical practices. Our models showed relatively lower performances than previous studies, but our methodology is promising in that we incorporated short-term or cross-sectional factors that could be accessible and readily measured in clinical settings, on top of the genetic factors. Despite the acceptable accuracy and sensitivity, however, there is much room for improvement regarding specificity and negative predictive value of the model.

This study has some limitations. We used European-ancestry based GWAS for estimating GPS for ethnically diverse samples. The transferability of our findings across different ethnic groups will be improved if better GPS methodology for multiethnic individuals is developed, which is currently an active area of research⁷⁵. Also, because of the scarcity of committed suicide in preadolescent youth, we had to use suicidal ideation and attempts as a proxy phenotype of suicidal behavior^{29,76}. It should be noted that not all suicidal ideation or attempts lead to committed suicide. We also acknowledge the potential selection bias from missing outcome/covariate data from the participants.

Our results highlight the importance and potential utility of the GPS approach in early screening which can identify a young population at high risk of suicidal behaviors. These findings also motivate the further development of effective screening methods and intervention strategies for youth at suicidal risk.

Declarations

Author contributions

Study concept and design: Y.J., J.C.

Acquisition, analysis, or interpretation of data: Y.J., H.K., S.M., H.W.

Drafting of the manuscript: Y.J., S.M., H.W., J.C.

Critical revision of the manuscript for important intellectual content: Y.J., J.K., I.C., J.C., W.-Y.A.

Statistical analysis: Y.J., S.M., H.W.

Obtained funding: I.C., J.C.

Study supervision: J.C.

Data and Code Availability

Codes and data are freely available for reproducibility

(<https://github.com/Transconnectome/ConnectomeLab/blob/master/suicide>).

Competing Interest Disclosures

None of the authors have significant competing financial, professional, or personal interests that might have influenced the performance or presentation of the work described in the manuscript.

Acknowledgement

This work was supported by the New Faculty Startup Fund from Seoul National University and Research grant from the Center for Happiness via the Center for Social Sciences at Seoul National University at Seoul National University.

References

1. Cunningham, R. M., Walton, M. A. & Carter, P. M. The Major Causes of Death in Children and Adolescents in the United States. *New England Journal of Medicine* vol. 379 2468–2475 (2018).
2. Website. National Institute of Mental Health. Mental health information. Statistics. Suicide. <https://www.nimh.nih.gov/health/statistics/suicide.shtml>.
3. Belsher, B. E. et al. Prediction Models for Suicide Attempts and Deaths: A Systematic Review and Simulation. *JAMA Psychiatry* 76, 642–651 (2019).
4. Brent, D. A., Bridge, J., Johnson, B. A. & Connolly, J. Suicidal behavior runs in families. A controlled family study of adolescent suicide victims. *Arch. Gen. Psychiatry* 53, 1145–1152 (1996).
5. Roy, A. & Segal, N. L. Suicidal behavior in twins: a replication. *J. Affect. Disord.* 66, 71–74 (2001).
6. Voracek, M. & Loibl, L. M. Genetics of suicide: a systematic review of twin studies. *Wien. Klin. Wochenschr.* 119, 463–475 (2007).
7. Mann, J. J., Waternaux, C., Haas, G. L. & Malone, K. M. Toward a clinical model of suicidal behavior in psychiatric patients. *Am. J. Psychiatry* 156, 181–189 (1999).
8. Sokolowski, M., Wasserman, J. & Wasserman, D. Polygenic associations of neurodevelopmental genes in suicide attempt. *Mol. Psychiatry* 21, 1381–1390 (2016).
9. Mullins, N. et al. Genetic relationships between suicide attempts, suicidal ideation and major psychiatric disorders: a genome-wide association and polygenic scoring study. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 165B, 428–437 (2014).
10. Docherty, A. R. et al. Genome-Wide Association Study of Suicide Death and Polygenic Prediction of Clinical Antecedents. *American Journal of Psychiatry* vol. 177 917–927 (2020).
11. Galfalvy, H. et al. A pilot genome wide association and gene expression array study of suicide with and without major depression. *World J. Biol. Psychiatry* 14, 574–582 (2013).
12. Willour, V. L. et al. A genome-wide association study of attempted suicide. *Molecular Psychiatry* vol. 17 433–444 (2012).
13. Perlis, R. H. et al. Genome-Wide Association Study of Suicide Attempts in Mood Disorder Patients. *American Journal of Psychiatry* vol. 167 1499–1507 (2010).
14. Mullins, N. et al. GWAS of Suicide Attempt in Psychiatric Disorders and Association With Major Depression Polygenic Risk Scores. *Am. J. Psychiatry* 176, 651–660 (2019).
15. Otsuka, I. et al. Genome-wide association studies identify polygenic effects for completed suicide in the Japanese population. *Neuropsychopharmacology* 44, 2119–2124 (2019).

16. Laje, G. et al. Genome-wide association study of suicidal ideation emerging during citalopram treatment of depressed outpatients. *Pharmacogenetics and Genomics* vol. 19 666–674 (2009).
17. Perroud, N. et al. Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project. *Pharmacogenomics J.* 12, 68–77 (2012).
18. Consortium, T. I. S. & The International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* vol. 460 748–752 (2009).
19. Consortium, S. W. G. of T. P. G. & Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* vol. 511 421–427 (2014).
20. Wray, N. R., Goddard, M. E. & Visscher, P. M. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Research* vol. 17 1520–1528 (2007).
21. Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet.* 9, e1003348 (2013).
22. Hunter, D. J. & Longo, D. L. The Precision of Evidence Needed to Practice ‘Precision Medicine’. *New England Journal of Medicine* vol. 380 2472–2474 (2019).
23. Sugrue, L. P. & Desikan, R. S. What Are Polygenic Scores and Why Are They Important? *JAMA* 321, 1820–1821 (2019).
24. Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* 19, 581–590 (2018).
25. Turecki, G., Ota, V. K., Belangero, S. I., Jackowski, A. & Kaufman, J. Early life adversity, genomic plasticity, and psychopathology. *Lancet Psychiatry* 1, 461–466 (2014).
26. Turecki, G. & Brent, D. A. Suicide and suicidal behaviour. *The Lancet* vol. 387 1227–1239 (2016).
27. Lewis, C. R. & Foster Olive, M. Early-life stress interactions with the epigenome. *Behavioural Pharmacology* vol. 25 341–351 (2014).
28. Cruceanu, C., Matosin, N. & Binder, E. B. Interactions of early-life stress with the genome and epigenome: from prenatal stress to psychiatric disorders. *Current Opinion in Behavioral Sciences* vol. 14 167–171 (2017).
29. Law, K. C., Jin, H. M. & Anestis, M. D. The intensity of suicidal ideation at the worst point and its association with suicide attempts. *Psychiatry Res.* 269, 524–528 (2018).
30. Conomos, M. P., Miller, M. B. & Thornton, T. A. Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. *Genet. Epidemiol.* 39, 276–293

(2015).

31. Conomos, M. P., Reiner, A. P., Weir, B. S. & Thornton, T. A. Model-free Estimation of Recent Genetic Relatedness. *Am. J. Hum. Genet.* 98, 127–148 (2016).
32. Pompili, M. et al. The Interplay Between Suicide Risk, Cognitive Vulnerability, Subjective Happiness and Depression Among Students. *Current Psychology* vol. 35 450–458 (2016).
33. Hsu, C.-Y., Chang, S.-S. & Yip, P. S. F. Subjective wellbeing, suicide and socioeconomic factors: an ecological analysis in Hong Kong. *Epidemiology and Psychiatric Sciences* vol. 28 112–130 (2019).
34. Hamilton, J. L. & Buysse, D. J. Reducing Suicidality Through Insomnia Treatment: Critical Next Steps in Suicide Prevention. *American Journal of Psychiatry* vol. 176 897–899 (2019).
35. Gabrielle A. Carlson, D. P. C. Suicidal Behavior and Depression in Children and Adolescents. *J. Am. Acad. Child Psychiatry* 21, 361–368 (1982).
36. King, R. A. et al. Psychosocial and risk behavior correlates of youth suicide attempts and suicidal ideation. *J. Am. Acad. Child Adolesc. Psychiatry* 40, 837–846 (2001).
37. Keilp, J. G. et al. Neuropsychological dysfunction in depressed suicide attempters. *Am. J. Psychiatry* 158, 735–741 (2001).
38. Rosoff, D. B., Kaminsky, Z. A., McIntosh, A. M., Davey Smith, G. & Lohoff, F. W. Educational attainment reduces the risk of suicide attempt among individuals with and without psychiatric disorders independent of cognition: a bidirectional and multivariable Mendelian randomization study with more than 815,000 participants. *Transl. Psychiatry* 10, 388 (2020).
39. Becker, S. P., Dvorsky, M. R., Holdaway, A. S. & Luebbe, A. M. Sleep problems and suicidal behaviors in college students. *J. Psychiatr. Res.* 99, 122–128 (2018).
40. Gorday, J. Y., Rogers, M. L. & Joiner, T. E. Examining characteristics of worry in relation to depression, anxiety, and suicidal ideation and attempts. *J. Psychiatr. Res.* 107, 97–103 (2018).
41. Price, C., Hemmingsson, T., Lewis, G., Zammit, S. & Allebeck, P. Cannabis and suicide: longitudinal study. *Br. J. Psychiatry* 195, 492–497 (2009).
42. Orri, M. et al. A genetically informed study on the association of cannabis, alcohol, and tobacco smoking with suicide attempt. *Molecular Psychiatry* (2020) doi:10.1038/s41380-020-0785-6.
43. Borges, G. et al. A meta-analysis of acute use of alcohol and the risk of suicide attempt. *Psychol. Med.* 47, 949–957 (2017).
44. Green, M., Turner, S. & Sareen, J. Smoking and suicide: biological and social evidence and causal mechanisms. *J. Epidemiol. Community Health* 71, 839–840 (2017).

45. Beauchaine, T. P., Ben-David, I. & Bos, M. ADHD, financial distress, and suicide in adulthood: A population study. *Sci Adv* 6, (2020).
46. Cassidy, S. & Rodgers, J. Understanding and prevention of suicide in autism. *Lancet Psychiatry* 4, e11 (2017).
47. Angst, J., Angst, F. & Stassen, H. H. Suicide risk in patients with major depressive disorder. *J. Clin. Psychiatry* 60 Suppl 2, 57–62; discussion 75–6, 113–6 (1999).
48. Palmer, B. A., Shane Pankratz, V. & Bostwick, J. M. The Lifetime Risk of Suicide in Schizophrenia. *Archives of General Psychiatry* vol. 62 247 (2005).
49. Hawton, K., Sutton, L., Haw, C., Sinclair, J. & Harriss, L. Suicide and attempted suicide in bipolar disorder: a systematic review of risk factors. *J. Clin. Psychiatry* 66, 693–704 (2005).
50. Gradus, J. L. et al. Posttraumatic Stress Disorder and Completed Suicide. *American Journal of Epidemiology* vol. 171 721–727 (2010).
51. Jansen, I. E. et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* 51, 404–413 (2019).
52. Choi, S. W., Mak, T. S.-H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* 15, 2759–2772 (2020).
53. Raj, A., Stephens, M. & Pritchard, J. K. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589 (2014).
54. Kaufman, J. et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* 36, 980–988 (1997).
55. Kaufman, J., Townsend, L. D. & Kobak, K. The Computerized Kiddie Schedule for Affective Disorders and Schizophrenia (KSADS): Development and Administration Guidelines. *Journal of the American Academy of Child & Adolescent Psychiatry* vol. 56 S357 (2017).
56. Luciana, M. et al. Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. *Dev. Cogn. Neurosci.* 32, 67–79 (2018).
57. Laan, M. J. van der, van der Laan, M. J., Polley, E. C. & Hubbard, A. E. Super Learner. *Statistical Applications in Genetics and Molecular Biology* vol. 6 (2007).
58. H2O Driverless AI - Open Source Leader in AI and ML. <https://www.h2o.ai/products/h2o-driverless-ai/>.

59. Chronis-Tuscano, A. et al. Very early predictors of adolescent depression and suicide attempts in children with attention-deficit/hyperactivity disorder. *Arch. Gen. Psychiatry* 67, 1044–1051 (2010).
60. McFadden, D. Conditional Logit Analysis of Qualitative Choice Behavior. (1973).
61. Campos, A. I. et al. Genetic aetiology of self-harm ideation and behaviour. *Sci. Rep.* 10, 9713 (2020).
62. Levey, D. F. et al. Genetic associations with suicide attempt severity and genetic overlap with major depression. *Transl. Psychiatry* 9, 22 (2019).
63. Shen, H. et al. Polygenic prediction and GWAS of depression, PTSD, and suicidal ideation/self-harm in a Peruvian cohort. *Neuropsychopharmacology* vol. 45 1595–1602 (2020).
64. Howe, S. J., Hewitt, K., Baraskewich, J., Cassidy, S. & McMorris, C. A. Suicidality Among Children and Youth With and Without Autism Spectrum Disorder: A Systematic Review of Existing Risk Assessment Tools. *J. Autism Dev. Disord.* 50, 3462–3476 (2020).
65. Chen, Y.-Y., Chen, Y.-L. & Gau, S. S.-F. Suicidality in Children with Elevated Autistic Traits. *Autism Res.* 13, 1811–1821 (2020).
66. Mayes, S. D., Gorman, A. A., Hillwig-Garcia, J. & Syed, E. Suicide ideation and attempts in children with autism. *Research in Autism Spectrum Disorders* vol. 7 109–119 (2013).
67. McDonnell, C. G. et al. An Exploratory Analysis of Predictors of Youth Suicide-Related Behaviors in Autism Spectrum Disorder: Implications for Prevention Science. *J. Autism Dev. Disord.* 50, 3531–3544 (2020).
68. Hirvikoski, T. et al. Individual risk and familial liability for suicide attempt and suicide in autism: a population-based study. *Psychol. Med.* 50, 1463–1474 (2020).
69. Ljung, T., Chen, Q., Lichtenstein, P. & Larsson, H. Common etiological factors of attention-deficit/hyperactivity disorder and suicidal behavior: a population-based study in Sweden. *JAMA Psychiatry* 71, 958–964 (2014).
70. Borges, G. et al. Twelve-month prevalence of and risk factors for suicide attempts in the World Health Organization World Mental Health Surveys. *J. Clin. Psychiatry* 71, 1617–1628 (2010).
71. Borges, G. et al. A risk index for 12-month suicide attempts in the National Comorbidity Survey Replication (NCS-R). *Psychol. Med.* 36, 1747–1757 (2006).
72. Mann, J. J. et al. Classification Trees Distinguish Suicide Attempters in Major Psychiatric Disorders. *The Journal of Clinical Psychiatry* vol. 69 23–31 (2008).
73. Glenn, C. R. & Nock, M. K. Improving the short-term prediction of suicidal behavior. *Am. J. Prev. Med.* 47, S176–80 (2014).

74. Huang, H. et al. Improving Polygenic Prediction in Ancestrally Diverse Populations. doi:10.21203/rs.3.rs-133290/v1.

75. Pigeon, W. R., Bishop, T. M. & Titus, C. E. The Relationship Between Sleep Disturbance, Suicidal Ideation, Suicide Attempts, and Suicide Among Adults: A Systematic Review. *Psychiatric Annals* vol. 46 177–186 (2016).

Tables

Table 1. Sociodemographic and clinical outcomes of the study participants from the Adolescent Brain and Cognitive Development (ABCD) study. The completed dataset of phenotypic outcomes, genotype, and covariates data available for 7,206 preadolescent U.S. children.

	Youth with suicidality (n=1,020)		Control Youth (N=6,186)	
	Count or Mean	Percentage (%) or SD	Count or Mean	Percentage (%) or SD
Female	397	38.9%	3,001	48.5%
Average Age (in months)	118.8	7.54	119.0	7.45
Average Income (bracket in 1~10 scales)	7.07	2.38	7.26	2.41
Marital status of the family (Currently married)	733	63.1%	4,987	70.7%
Average Mother's EA	16.62	2.63	16.61	2.78

Table 2. Significant associations of GPSs of 25 psychiatric and common traits with youth suicidality, surviving FDR significance < 0.05.

(a) Analysis of Entire multiethnic participants (n=7,206)

Outcome	GPS predictor	OR	95% CI (lower)	95% CI (upper)	P-value	N of cases
Suicidal Ideation (Active + Passive)	ADHD	1.13	1.06	1.21	3.48E-04	1014
Suicidality(all)	ADHD	1.13	1.05	1.21	4.64E-04	1020
Active Suicidal Ideation	ADHD	1.17	1.07	1.28	4.71E-04	540
Passive Suicidal Ideation	ADHD	1.13	1.05	1.21	1.29E-03	832
Passive Suicidal Ideation	GENERALHAPPINESS	0.89	0.83	0.96	1.43E-03	832
Suicidal Ideation (Active + Passive)	GENERALHAPPINESS	0.90	0.84	0.96	1.47E-03	1014
Active Suicidal Ideation	ASD	1.15	1.05	1.26	1.81E-03	540
Suicidality(all)	GENERALHAPPINESS	0.90	0.84	0.96	2.41E-03	1020

(b) Analysis of European participants (n=5,749)

Outcome	GPS predictor	OR	95% CI (lower)	95% CI (upper)	P-value	N of cases
Suicidal Ideation (Active + Passive)	ADHD	1.16	1.08	1.25	5.60E-05	796
Suicidality(all)	ADHD	1.16	1.08	1.25	6.60E-05	798
Passive Suicidal Ideation	ADHD	1.17	1.08	1.27	1.39E-04	663
Active Suicidal Ideation	ADHD	1.20	1.08	1.32	4.22E-04	409
Passive Suicidal Ideation	GENERALHAPPINESS	0.87	0.81	0.95	8.61E-04	663
Suicidality(all)	SMOKER	1.13	1.05	1.22	1.08E-03	798
Suicidal Ideation (Active + Passive)	SMOKER	1.13	1.05	1.22	1.18E-03	796
Active Suicidal Ideation	ASD	1.18	1.07	1.3	1.26E-03	409
Suicidal Ideation (Active + Passive)	GENERALHAPPINESS	0.89	0.83	0.96	1.77E-03	796
Suicidality(all)	GENERALHAPPINESS	0.89	0.83	0.96	1.97E-03	798
Suicidality(all)	MDD	1.12	1.04	1.2	2.65E-03	798
Suicidal Ideation (Active + Passive)	MDD	1.12	1.04	1.2	2.78E-03	796
Passive Suicidal Ideation	RISK4PC	1.13	1.04	1.22	3.43E-03	663
Suicide Attempt	SCZ	1.51	1.14	1.99	3.55E-03	48
Suicidality(all)	RISK4PC	1.11	1.03	1.2	5.36E-03	798
Passive Suicidal Ideation	SMOKER	1.12	1.03	1.21	6.12E-03	663
Suicidal Ideation (Active + Passive)	RISK4PC	1.11	1.03	1.19	6.27E-03	796
Passive Suicidal Ideation	MDD	1.12	1.03	1.21	6.68E-03	663

Figures

Figure 1

Consort flow diagram for the study. The study initially assessed 11,869 preadolescent children ages of 9 to 10 years old recruited from the Adolescent Brain and Cognitive Development (ABCD) study. The completed dataset of phenotypic outcomes, genotype, and covariates data available for 7,206 preadolescent children and used for the analysis.

Figure 2

Manhattan plot showing the associations of 25 genome-wide polygenic scores (GPS) of cognitive, psychological, personality traits, and psychiatric disorders with childhood suicidality. The PheWAS analysis of the GPSs was performed on the (a) 7,202 multiethnic individuals and (b) 5,749 European-ancestry individuals at age 9-10. Red line indicates FDR corrected p-value of 0.05 and target GPSs are color-coded. Each point represents the effect direction (positive/negative) and effect size of each GPS-phenotype association. An upward arrow indicates a positive effect size of the association.

Figure 3

Prediction performance of the machine learning models based on GPS and cognitive, psychological, behavioral, environmental, familial variables. Receiver-Operator Characteristics (ROC) curves of the models predicting suicidal ideation (A) and suicide attempt (B) using CatBoost classifier. In predicting suicidal ideation, 6,106 subjects (European-ancestry=4,863) are used for training and 366 subjects (European-ancestry=294) for testing. On the suicidal attempt, 6,442 subjects (European-ancestry=5,103) are used for training and 25 subjects (European-ancestry=22) for testing. We also tested the prediction model only with European-ancestry (EA) and compared it with the results of entire ancestry.