



OPEN

Machine learning of clinical and neural data predicts future homicide in high-risk youth

Samantha N. Rodriguez^{1,2}, Aparna R. Gullapalli¹, David D. Stephenson¹, J. Michael Maurer¹, Keith A. Harenski¹, Carla L. Harenski¹ & Kent A. Kiehl^{1,2}✉

Homicide is one of the most severe forms of violence and is the single most costly crime in the United States. Despite its high financial cost and incalculable emotional toll, we know very little about predicting who will commit such crimes. Here we followed a large ($n=202$) sample of incarcerated youth for 16 years after release from a maximum-security juvenile correctional facility. Official records and clinical interviews indicated $n=35$ committed a homicide during this follow-up period. Clinical (psychopathic traits, age of first arrest) and neural variables (gray matter volume of the amygdala and temporal pole) collected at baseline while boys were housed at the correctional facility significantly differed between homicide and non-homicide ($n=167$) groups. Classification through machine learning models using these clinical and neural data predicted which formerly incarcerated youth committed a future homicide as adults with high accuracy. The model which included a priori clinical (psychopathic traits) and neural variables (bilateral amygdala, insula, parahippocampal gyrus, middle temporal pole, superior temporal pole, and orbitofrontal cortex) achieved top performance with 76% accuracy (sensitivity = 86%; specificity = 75%). The implications of these results are discussed as they relate to intervention and prevention efforts.

Keywords Machine learning, Homicide, Prediction, Juvenile offenders, Neuroimaging, Forensic psychology

Crime is a major public health issue in the United States. Societal estimates of the cost of crime exceed \$4 trillion annually¹. Homicide, the death of an individual by means of violence committed by another person, is the single most expensive offense-costing 20 times more than any other crime². The emotional cost associated with loss of life is incalculable. The majority of homicides are committed by individuals who have been previously arrested³. Further, juvenile antisocial behavior is a significant predictor of persistent and severe criminal behavior occurring during adulthood⁴. Despite this, underlying clinical features and neural circuits associated with homicide remain poorly understood.

Postdictive studies have investigated characteristics and past experiences of juveniles who already committed homicide. In one study, youth convicted of homicide, compared to youth convicted of burglary, were more likely to have authoritarian parents, parental arrest record, and extensive problems at school⁵. Another study comparing homicidal youth to non-homicidal youth found lower IQ and higher rates of exposure to violence in those who previously committed homicide⁷. In a study comparing youth with history of homicide with youth who had committed non-homicidal violent acts, homicidal youth reported higher frequency of substance use disorders and greater availability of firearms⁸. Finally, research from our group suggests homicidal youth are characterized by higher rates of substance use, post-traumatic stress disorder, and psychopathic traits compared to non-homicidal violent and non-violent incarcerated youth⁶.

The first prospective studies identifying variables associated with future homicide were conducted using the Pittsburgh Youth Study sample, a multi-cohort sample representative of school-age children. Researchers found having a disruptive behavioral disorder, serious delinquency in self and peers, cruelty towards people, school delinquency, and positive attitudes towards drugs and delinquency increased odds ratio of committing a future homicide from 1.9 to 4.9⁹. They also found a cumulative effect of risk factors, such that having four or more violence risk factors (e.g., behavioral, attitude, cognition, psychiatric, offending, birth, family, peer, and school factors) presented increased risk of future homicidal behavior. Another study re-examining the Pittsburgh Youth Study sample reported the individuals characterized by a higher number of early risk factors were associated with an increased probability of later conviction for violent offences, including homicide¹⁰.

¹The Mind Research Network, 1101 Yale Blvd NE, Albuquerque, NM 87106, United States of America. ²Department of Psychology, University of New Mexico, 2001 Redondo S. Dr, Albuquerque, NM 87106, United States of America. ✉email: kkiehl@unm.edu

Early risk factors included experiencing physical abuse, parental stress, associating with delinquent peers, and low school motivation¹⁰. These studies highlight social and clinical variables that can promote risk for future violence, including homicide. However, none of the latter studies have examined the neural circuits that might differ between youth who do and do not commit homicide.

Our research group has previously identified neural differences between incarcerated youth who have and have not previously committed a homicide⁶. Incarcerated boys who previously committed a homicide were characterized by reduced gray matter volume in the medial and lateral temporal lobe compared to incarcerated boys with no prior history of homicide. However, no study to date has investigated whether clinical and neuroimaging data collected during adolescence from incarcerated youth can predict future instances of homicide. Here, we present the first longitudinal study using clinical and neuroimaging variables collected from incarcerated adolescents to prospectively predict future homicides committed during adulthood. We hypothesized that boys who committed a future homicide would present with higher psychopathic traits than their incarcerated peers who would not go on to commit homicide (see¹¹ for meta-analysis). Early age of engaging in antisocial behavior has long been theorized to characterize youth who have extensive criminal careers continuing into adulthood^{4,12}. Given this, we also hypothesized earlier age of onset of antisocial behavior in boys who would commit a future homicide. Additionally, based on our prior research⁶, we hypothesized that reduced regional gray matter volume in the bilateral amygdala, insula, parahippocampal gyrus, middle and superior temporal pole, and orbitofrontal cortex collected during adolescence would predict which boys committed a future homicide during adulthood. Machine learning classification analyses, using support-vector machine (SVM), were performed to identify clinical and neural variables that differentiated individuals who did and did not commit a future homicide. The goal of this work is to identify both clinical and neurobiological targets for intervention, such that tailored treatment can be implemented early in life to help prevent future instances of severe violent behavior.

Results

Group Differences

Participants in the current study are from the SouthWest Advanced Neuroimaging Cohort-Youth (SWANC-Y; funded by NIMH R01MH071896-0: PI: Kiehl). Comprehensive criminal records review and self-reported incidents of homicide gathered from clinical interviews were used to determine group assignment. Of the $n = 202$ incarcerated boys studied here, $n = 35$ committed a homicide after their release from the juvenile correctional facility (i.e., "Homicide" (H) group). The remaining $n = 167$ boys who showed no evidence of homicide crimes before or during the follow-up period were assigned to the "No-Homicide" (No-H) group. See "Materials and Methods" for additional information on group classification strategy. Consistent with our hypotheses, individuals in the H group scored significantly higher on baseline Hare Psychopathy Checklist: Youth Version (PCL: YV¹³) Total scores ($M = 25.9$, $SD = 4.8$) as compared to No-H participants ($M = 22.4$, $SD = 5.5$), $t(200) = 3.60$, $p < 0.001$. Participants in the H group also scored higher on PCL: YV Factor 1 scores ($M = 7.7$, $SD = 2.3$) than No-H ($M = 6.1$, $SD = 2.8$) participants, $t(200) = 3.10$, $p = 0.002$. Additionally, participants in the H group scored significantly higher on PCL: YV Factor 2 scores ($M = 15.8$, $SD = 2.4$) than participants in the No-H group ($M = 14.1$, $SD = 3.0$), $t(200) = 3.32$, $p = 0.001$. Finally, H participants were significantly younger ($M = 11.8$, $SD = 2.5$) than No-H participants ($M = 12.6$, $SD = 2.1$) at the time of their first arrest, $t(200) = -2.14$, $p = 0.033$. Groups did not significantly differ on any of the remaining 23 clinical variables investigated (see Supplementary Tables S3 and S4 for descriptive statistics, and Supplementary Tables S5 and S6 for results). Supplementary Tables S7 and S8 display exploratory t -test results between H and No-H groups on PCL: YV Facet and Item scores. Further, approximately 30% of the youth ($n = 11$) in the H group had committed a homicide during adolescence. In order to examine the role of past behavior in predicting future behavior, we conducted t -test analyses excluding these participants from the H group. The Future Homicide-Only (FH-only) group was then compared to No-H group (see Supplementary Tables S9 and S10). Significant differences in clinical variables between the original H-group versus the No-H group were also observed in the FH-Only versus the No-H group analyses.

As predicted, H participants exhibited reduced gray matter volume within a priori ROIs, including the bilateral amygdala, bilateral middle temporal pole, and right superior temporal pole compared to No-H participants. There were no significant differences in the remaining seven ROIs. Whole-brain voxel-based morphometry (VBM) analyses and group difference statistics are presented in Fig. 1; Table 1. Results of a VBM analysis comparing FH-Only and No-H groups can be found in Supplementary Fig. S3. These results are similar to, albeit less widespread, than the original VBM analysis comparing H and No-H participants.

Machine learning: support-vector machine (SVM)

To test the predictive capability of the imaging and behavioral variables while making no assumptions on their relationships or the underlying distributions of these variables (unlike a regression) a binary classifier using machine learning was trained to predict H and No-H participants. Machine learning classification was used to develop models that identified the best possible combination of variables that could accurately predict between H and No-H participants. Due to the lack of an existing separate forensic sample to test the generalizability of the machine learning model, of the $n = 202$ individuals included in our study, 20% of participants in both H and No-H groups were held out as an out-of-sample test set on which the trained machine learning model would be evaluated for performance. The choice of 20% was guided by having enough of the Homicide group in the test dataset while retaining a sufficient number in the model building dataset to build a generalizable model. The remaining 80% of participants were used as the data sample to tune a linear SVM machine learning model using a stratified, k -fold cross-validated approach with $k = 5$ to ensure enough of the rare H samples were retained in the test sample of each resampling fold thereby balancing the tradeoff between bias and variance in the trained model, given the low base rate of H in the overall sample^{14–16}. Repeated cross validation and

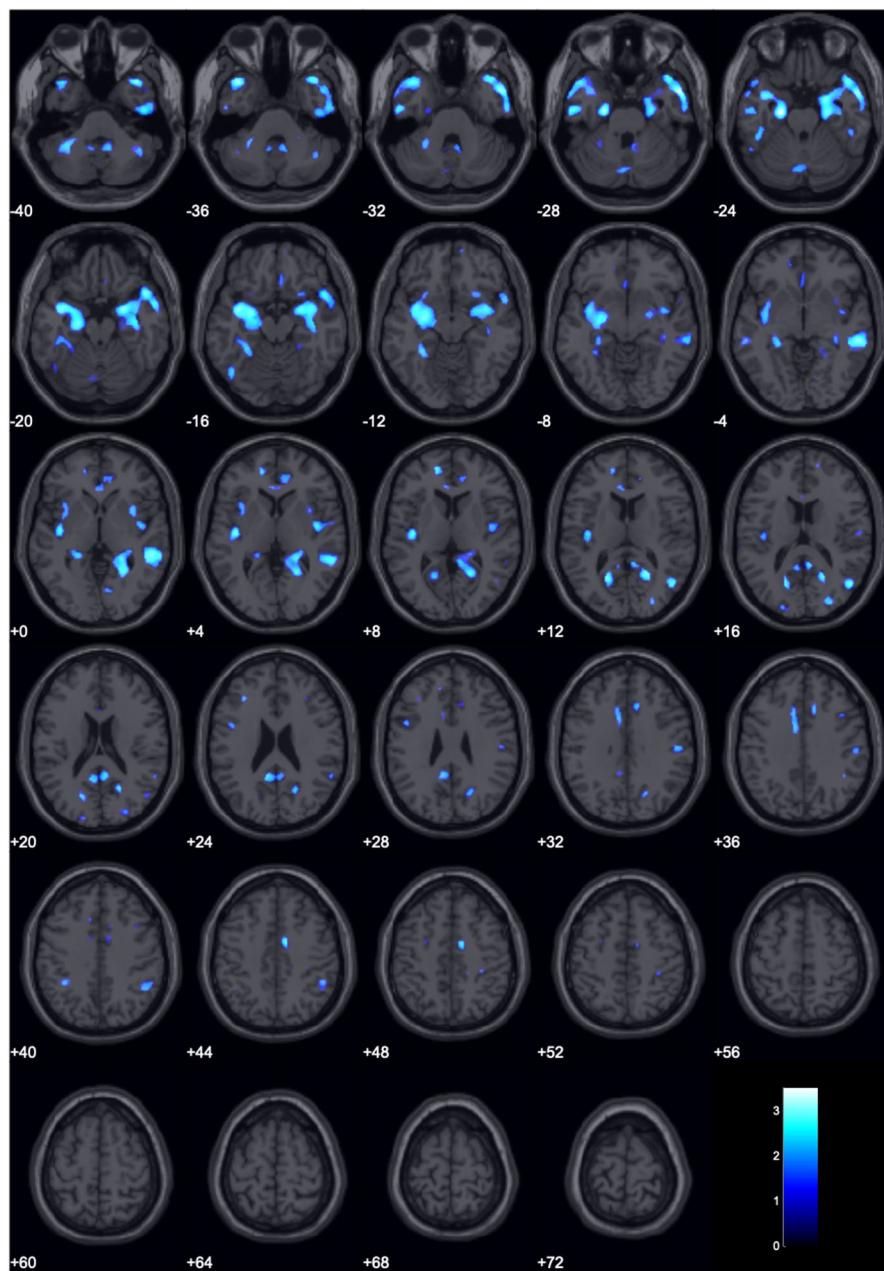


Fig. 1. Voxel-based morphometry (VBM) results comparing gray matter volume (GMV) between Homicide (H; $n=35$) and No-Homicide (No-H; $n=167$) participants. Covariates included in VBM analyses were age at MRI scan and total brain volume. Blue indicates t -values for regions in which the H group had lower GMV compared to the No-H group, with lighter blue/white indicating larger differences. Results presented at $p < 0.05$, uncorrected.

stratified cross validation approaches¹⁶ were implemented during model training to minimize variance and bias. Keeping with best practices of handling prediction in a small data set with a relatively rare occurrence event, a cost-sensitive SVM kernel was used for learning the model, the classifiers were evaluated on the balanced accuracy metric and the threshold for classification was tuned on the final model before applying to the test data¹⁷ (See Machine Learning subsection of Methods for more detailed methodology). The best performing model (76% overall accuracy, 86% sensitivity, and 74% specificity) consisted of gray matter volume estimates for all 12 a priori ROIs, total brain volume, age at scan, and PCL: YV factor scores. Variables included in all models are listed in Supplementary Table S11. Model statistics are displayed in Table 2 and Supplementary Table S12. Feature importance analyses were also run to determine the extent to which each variable contributed to the classification rate for each respective model. See Fig. 2 for feature importance values for the best performing model, Fig. 3 for feature importance values for all models run, and Supplementary Fig. S1 for a comparison

Region	x	y	z	t (200)	Cohen's d	p-FWE	p-FDR	p-un
L. Amygdala	-27	-6	-14	3.19	0.45	0.010	0.012	0.0004
R. Amygdala	32	-2	-15	3.15	0.45	0.013	0.018	0.0005
L. Insula	-36	0	-15	2.95	0.42	0.092	0.176	0.0009
R. Insula	45	-8	3	2.51	0.35	0.211	0.232	0.0032
L. Parahippocampal gyrus	-23	-14	-26	2.70	0.38	0.101	0.154	0.0019
R. Parahippocampal gyrus	29	-12	-26	2.54	0.36	0.163	0.202	0.0030
L. Middle temporal pole	-38	21	-36	3.20	0.45	0.031	0.041	0.0004
R. Middle temporal pole	38	18	-42	3.56	0.50	0.016	0.012	0.0001
L. Superior temporal pole	-39	21	-33	3.04	0.43	0.058	0.254	0.0007
R. Superior temporal pole	50	18	-21	3.18	0.45	0.049	0.025	0.0004
L. Orbitofrontal cortex	-32	38	-8	1.71	0.24	0.424	0.495	0.0217
R. Orbitofrontal cortex	48	118	-15	2.33	0.33	0.287	0.422	0.0052

Table 1. Neuroanatomical differences between homicide (H; $n=35$) and No-Homicide (No-H; $n=167$) participants. Reductions in Gray matter volume in a priori regions of interest in H participants as compared to No-H participants; $t=t$ -value; p -FWE = p -value corrected for family-wise error rate; p -FDR = p -value corrected for false discovery rate; p -un = uncorrected, Raw p -value; bolded p -values indicate results which survived FWE-correction or FDR-correction in the corresponding column.

Model	Model performance (%) overall/sensitivity/specificity	ROC AUC (%)
A priori model with neural data + Age + PCL: YV scores	76/86/74	82
Clinical data only	68/14/80	65
Neural data only	71/71/71	80
Clinical + neural data	68/86/65	80

Table 2. Prediction results across machine learning models. Numbers in model performance column represent percentages of overall accuracy, sensitivity (i.e., true positive), and specificity (i.e., true negative) rates, respectively. Area under the receiver operating characteristic (ROC) curve results are presented in the last column. The ROC curve plots true positive rates against false positive rates for all decision thresholds. The ROC AUC computes the area under the ROC curve, with a higher value indicating better ability to predict homicide group members (positive) while minimizing false positives (1: perfect predictive model, 0.5: random guess model). A priori model is listed first. Remaining models were run for completeness. Clinical Data = Age, IQ, socioeconomic status (SES), number of substance use disorders (SUD), addiction severity index (ASI) substance use severity scores, Barrett impulsiveness scale (BIS-11) total scores, psychopathy checklist: youth version (PCL: YV) factor 1 and factor 2 scores, number of traumatic brain injuries (TBI), age of first arrest, trauma checklist 2.0 (TCL) factor 1 and factor 2 scores, number of prior felonies, probation/parole violations, and convictions, psychopathology assessed via kiddie schedule for affective disorders and schizophrenia (KSADS) (i.e., anxiety, depression, conduct disorder/ oppositional deviant disorder (CD/ODD), attention deficit/ hyperactivity disorder (ADHD), post-traumatic stress disorder (PTSD)), gang affiliation, parental incarceration, parental separation; neural Data = GMV within a priori ROIs including the left and right amygdala, left and right insula, left and right parahippocampal gyrus, left and right middle temporal pole, left and right superior temporal pole, left and right orbitofrontal cortex, and total brain volume (BV).

of feature importance values of all models. Receiver operating characteristic (ROC) curves for all models are displayed in Supplementary Fig. S2.

Discussion

The current study identified clinical and neural variables collected during adolescence that were predictive of future instances of homicide occurring during adulthood. A sample of $n=202$ high-risk adolescents incarcerated at a maximum-security juvenile correctional facility in the United States were followed for a period of 16 years post-release. We observed that adolescents from this sample who committed a future homicide as adults scored higher on baseline measures of psychopathic traits, had an earlier age of onset of antisocial behavior, and were characterized by reduced GMV in bilateral amygdala, middle temporal pole, and right superior temporal pole compared to adolescents who did not commit a future homicide. Support-vector machine learning was able to predict participants who did (sensitivity: 86%) and did not (specificity: 74%) commit a future homicide with 76% overall accuracy and 82% area under the receiver operating characteristic (ROC) curve. This high sensitivity rate demonstrates that we were able to successfully predict six out of the seven (86%) incarcerated adolescents in the hold-out test sample who committed a future homicide as adults using both clinical and neural features.

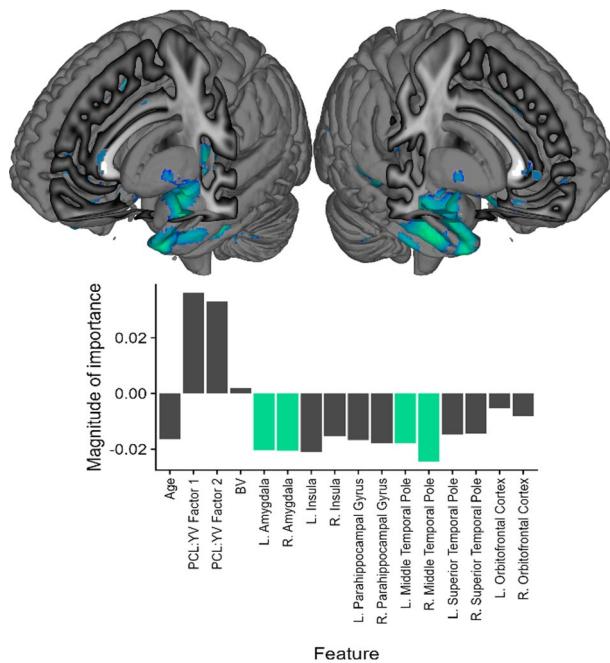


Fig. 2. 3D surface render of left and right sagittal orientations of the cortical and subcortical areas with Homicide < No-Homicide parametric overlay with respective feature importance. Left and right 3D surface renders of the cortical and subcortical areas of the brain in the sagittal anatomical orientation. Cut planes have been made in Montreal Neurological Institute (MNI) stereotaxic space at an intersection of $x = 5$, $y = -15$, and $z = -20$ to simultaneously show statistically significant cortical and subcortical effects. A parametric overlay of the Homicide < No-Homicide contrast map is rendered at a threshold of $p \leq 0.05$ and cluster extent threshold of $k = 10$ voxels on a gradient of blue to cyan indicating level of significance. Below the 3D surface render is the feature importance bar plot for the best performing classification model, with those neural features that significantly differed between groups visible in the above 3D brain render having their respective bars shaded in cyan. Taller bars indicate variables that contributed more strongly to the classification. Model displayed includes GMV within all 12 a priori ROIs, age at scan, BV, PCL: YV Factor 1, and PCL: YV Factor 2. PCL: YV Factor 1 and Factor 2, bilateral amygdala, left insula, and right middle temporal pole are the most strongly weighted variables contributing to group classification.

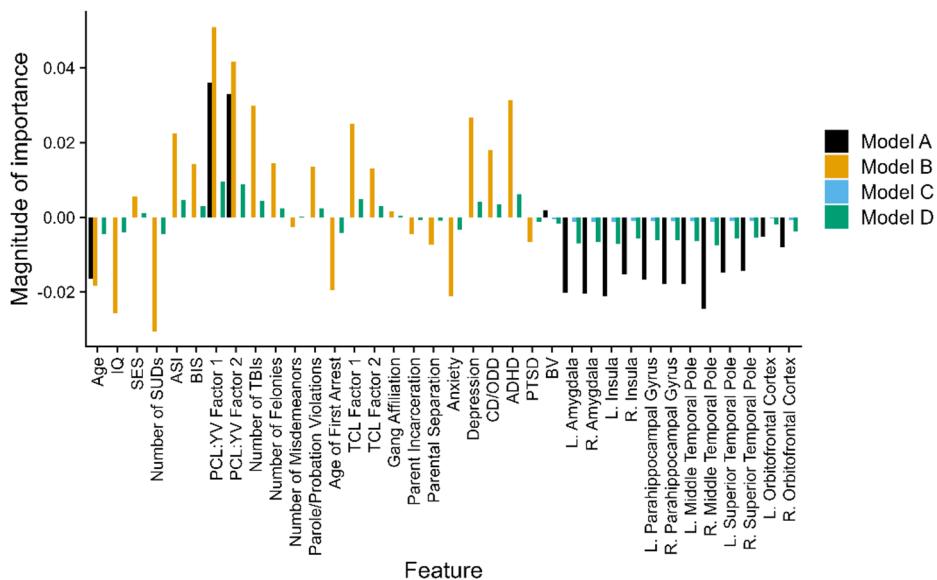


Fig. 3. Feature importance bars for all models. Bars represent the relative weight each variable held in group classification, with taller bars indicating that variable contributed more strongly to the classification. Model A = a priori model with neural data + Age + PCL: YV scores. Model B = clinical data only. Model C = neural data only. Model D = clinical + neural data.

Using a model that included both clinical (i.e., PCL: YV Factor 1 and Factor 2 scores, age at MRI scan) and neural data (i.e., 12 a priori ROI volume estimates and total BV), we were able to reliably differentiate between adolescents who did and did not commit a future homicide during adulthood. A model including age of first arrest in addition to these aforementioned variables was also run, and resulted in identical classification metrics. Age of first arrest was not included in final models presented here to reduce dimensionality. The 86% level of sensitivity remained consistent across model iterations and threshold tuning for models including both clinical and neural data (see Table 2 and Supplementary Table S12), demonstrating there is improved utility when incorporating both clinical measures and neuroanatomical metrics in predicting future homicidal behavior among high-risk youth. These classification results are similar to our previous report⁶ in postdictive classification of homicide offending during adolescence, where 81% overall accuracy, 81% sensitivity, and 80% specificity were found using a machine learning algorithm that included PCL: YV Factor 1 scores, total number of convictions, socioeconomic status, total brain volume, and GMV data from the OFC, cingulate cortex, and temporal pole. Importantly, the a priori model outlined in the current study was able to classify homicide offenders above and beyond just those who committed a violent offense. Specifically, classification into violent versus non-violent groups using this model was just above chance (see Supplementary Table S13). Further, $n = 15$ of the $n = 25$ correctly classified No-H participants in the hold-on sample had violent felony arrests in adulthood. Taken together, these studies demonstrate the applicability of machine learning in classifying ultra-high-risk adolescents.

It is important to note that not all models tested in the current study performed equally as well. Predicting future homicidal behavior using only clinical variables (i.e., psychopathic traits, history of trauma, previous offense history, etc.; for full list of variables included in all models see Supplementary Table S11) resulted in 65% ROC. When only including neural data in the classification model, model performance increased to 80% ROC, with 71% sensitivity and 71% specificity rates. While overall performance of this latter model is similar to models incorporating both clinical and neural data, the sensitivity rate is lower when only including neural data. These results support the notion that the inclusion of both clinical and neural data together serve as the best method for identifying at-risk adolescents most likely to commit a future homicide, while balancing overall model performance. Indeed, neural data may allow for the capture of latent traits not fully assessed using clinical instruments alone.

Our group difference results demonstrate that age of first arrest and psychopathic traits are particularly important in determining juveniles' risk for future homicide. It has been theorized that two subtypes of juvenile offenders exist: adolescent-limited and lifecourse persistent⁴. Adolescent-limited offenders engage in antisocial behavior only during childhood/adolescence and typically engage in lower rates of criminal offending upon entering adulthood. On the contrary, lifecourse persistent offenders continue to offend throughout the lifespan, with antisocial behavior routinely increasing in severity and frequency⁴. Lifecourse persistent offenders, compared to adolescent-limited, are characterized by earlier age of criminal onset⁴. The current results support this notion, as boys who would go on to commit a homicide in adulthood were significantly younger at the time of their first arrest compared to boys who would not commit a future homicide. This suggests boys included in our homicide group may be characterized as lifecourse-persistent offenders. Psychopathy scores may perhaps be even more discriminatory of who may commit a future homicide than age of first arrest. In this sample, boys in the H group scored significantly higher on PCL: YV Total, Factor 1, and Factor 2 scores. Feature importance models (Figs. 2 and 3, and Supplementary Fig. S1) displaying the relative predictive weight of each variable across all models show psychopathic traits are among the strongest predictors in all clinical-only and combined clinical and neural models. Psychopathic traits have also proven to have great utility in predicting future violence in other work^{18–21}.

In the present study, incarcerated adolescents who committed a future homicide during adulthood were characterized by reduced baseline gray matter volume in the bilateral amygdala, bilateral middle temporal pole, and right superior temporal pole compared to incarcerated adolescents who did not commit a future homicide. Further, these regions were among the strongest negatively weighted features in the classification model, indicating that reduction of volume in these regions contributed the most to group membership. Figure 3 and Fig. S1 demonstrate that when clinical and neural variables are included in the same model, neural data subsumes predictive information provided by the majority of clinical variables. Thus, reinforcing the importance of evaluating gray matter volume in regions identified here when assessing risk for homicide. Specifically, while the quality and breadth of information gained from clinical assessments varies by administrator and tool, neuroanatomical data is highly reliable and may be able to provide information on traits not directly assessed via clinical methods, such as is demonstrated here.

Reduced GMV within the amygdala has been previously shown to be predictive of both previous homicide in boys⁶ and future instances of violent felonies in adult men²². The amygdala has been previously implicated in processes including abnormal emotional learning and aversive conditioning in both psychopathy and antisocial behavior more generally²³. The temporal pole has been associated with emotional and social processing²⁴ and higher likelihood of rearrest²⁵. Studies have also found deficits in this region in those who score high in psychopathy²⁶. The current results demonstrate that incarcerated adolescents who would go on to commit a future homicide during adulthood are characterized by neuroanatomical abnormalities in regions contributing to social interactions and emotional responding. Additionally, GMV reductions in these regions have also been associated with childhood trauma exposure²⁷. It is important to consider, then, that neuroanatomic characteristics measured at baseline may in part be a consequence of earlier adverse experiences, including childhood maltreatment or trauma exposure. We note however, that the H and non-H groups did not significantly differ on an expert-rated assessment of trauma in the current study (see Supplementary Table S5). An additional VBM analysis was conducted including Trauma Checklist (TCL²⁸) as a covariate (Supplementary Fig. S4), which resulted in nearly identical findings as displayed in Fig. 1. Therefore, childhood maltreatment or exposure to trauma is not likely contributing significantly to the current results.

All incarcerated adolescents included in the SWANC-Y cohort are considered to be at a high risk for future criminal offending based on their previous antisocial behavior. However, being able to better understand the variables that place an adolescent at a more severe level of risk for future criminal behavior during adulthood, including homicide, is essential for improving the individual's quality of life, the safety and well-being of society overall, and reducing the cost of crime. A number of risk assessment tools are currently used to predict risk for future antisocial behavior in youth. Tools such as the Structured Assessment of Violence Risk in Youth (SAVRY)²⁹, Youth Assessment and Screening Instrument (YASI)³⁰, and Structured Assessment of Protective Factors – Youth Version (SAPROF-YV)³¹ are used in the justice system to provide insight on the likelihood of re-offending among juvenile offenders when making parole, sentencing, and treatment course decisions. Studies investigating the predictive ability of these tools find overall moderate predictive validity and have thus far been limited to predicting general or violent recidivism^{32–35}. To our knowledge, no study to date has examined the efficacy of such risk assessment tools in predicting future instances of homicide. Future studies may aim to compare the efficacy of the PCL: YV to other existing risk assessment tools in predicting future homicidal behavior.

Studies suggest that adolescence is a prime developmental window for intervention work^{36–38}. By determining risk level and providing the appropriate treatment during a crucial developmental period, reductions in future antisocial behavior have been observed in other samples of high-risk boys^{36–38}. For example, the Mendota Juvenile Treatment Center (MJTC) follows a clinical-correctional hybrid model³⁶ where high-risk youth are provided with intensive cognitive behavioral therapy, education, and ancillary services. MJTC uses positive reinforcement to reward good behavior and promotes attendance in treatment and programming with mental health professionals³⁶. Youth who completed programming at MJTC are 50% less likely to commit violent offenses upon release to the community than youth serving time in other correctional facilities. Indeed, in one study, none of the $n = 200$ MJTC treated youth committed a future homicide during a two-year follow-up window, whereas 10.6% of the untreated youth committed 16 total homicides³⁸. Neuroanatomical features like those identified in the current study may also represent key targets for specialized treatment modalities. Research suggests neuroplasticity, or the brain's ability to change over time, may be linked to behavioral changes resulting from psychotherapy^{39,40}. This, combined with the reduced prevalence of homicide in MJTC-treated youth, supports the idea that risk factors for future homicide identified in the current study (i.e., high psychopathy scores and reduced GMV in the amygdala and temporal pole) may be mitigated if appropriately addressed via early interventions.

We note that more research is needed before the methodologies presented here should be used for individual-subject prediction. Additionally, we acknowledge that MRI scans may not be readily available to all justice-involved youth. Nevertheless, our results are the first to provide support for the utility in combining clinical and neural data when investigating risk level for future homicide in adolescent offenders, up to 16-years after their release.

Limitations and Future Directions

There are a few limitations to note. Despite our best attempts to accurately assign group membership, it is possible that participants included in the No-H group committed a homicide following their release from the juvenile correctional facility and have been incorrectly assigned to the wrong group. While we used both official record ($n = 22$) and self-reported incidences ($n = 13$) of homicidal behavior, there may be participants included in the No-H group who have committed this behavior but were not caught and did not disclose this information to research staff. There may also be participants who have not yet committed a homicide and may go on to commit one in the future. We plan to conduct additional follow-up studies at longer timepoints to address this issue.

Our sample was unique in several ways. This project was designed to work with a very high-risk sample to help the judicial system make evidence-based decisions regarding both risk prediction and treatment programming. Our sample was drawn from the highest-security juvenile correctional center in the State of New Mexico. This is a state-run facility for serious adolescent offenders and our sample is likely to differ from secure county/city managed correctional facilities and certainly will differ from community-based samples. Our participants were only those research-seeking individuals housed at this juvenile correctional facility. Although most of the youth at this facility volunteered to participate in this project, our work may not generalize to the entire population of high-risk youth offenders (i.e., those who did not volunteer for research or were excluded [see exclusion criteria]). It is also possible that geospatial variables may influence results (i.e., state- or nation-wide homicide rate, firearm access), especially as it pertains to samples outside of the United States. Our sample was also restricted to boys. Thus, it is important for future research to examine other samples (e.g., incarcerated girls, adolescents housed at county or city-run facilities, and international samples) to test the generalizability of our findings.

Future studies should also aim to replicate and extend the methodology outlined here. Specifically, the current study identified variables predictive of any type of future homicide. Future studies should investigate specific homicidal subtypes, including sexual, multi-victim, or serial homicide. Additionally, while the best performing model had a high rate of specificity (i.e., true negatives), there were still boys included in the No-H group who were misclassified as high risk for future homicide. Given the practical implications of this work, more research is needed before the current methodologies can be used in practice. Finally, total or factor scores were included in models here. Future studies should investigate the utility of facet and/or individual item scores from various assessments in predicting future behavior. Thus, while the models explored here provide good groundwork for a psychobiological-based evaluation of risk factors associated with homicide, future studies are still needed before this or similar work should be implemented in informing risk level during probation, parole, and resentencing decisions.

Conclusion

The current study utilized machine learning methods to predict which juvenile offenders would and would not commit a homicide up to 16 years after their release from a juvenile correctional facility based on clinical and neuroanatomical data. Overall classification rates were high, with models achieving 76% overall accuracy (82% ROC), and sensitivity (predicting future homicide: 86%), and specificity (predicting non-homicide: 74%) rates. This study is the first to investigate the efficacy of both clinical and neuroimaging data collected from juvenile boys who are already deemed to be at a high risk for future antisocial behavior in predicting who will commit a future homicide during adulthood. This study serves as a promising foundation for future investigations into the validity of this approach in understanding and predicting risk for future offending, particularly homicide. These results also indicate psychopathic traits, supplemented by abnormalities in brain regions associated with emotional learning and prosocial behavior, may be key targets for specialized treatment approaches aimed at reducing risk for severe violent behavior.

Materials and Methods

Data Collection

Data for the current study was collected in two phases. First, data (e.g., clinical and neural variables) was originally collected from incarcerated male adolescents as part of the SouthWest Advanced Neuroimaging Cohort – Youth (SWANC-Y; R01 MH071896-01, PI: Kiehl) sample. Participants included in the SWANC-Y sample were recruited from a maximum-security juvenile correctional facility in the state of New Mexico from the years 2007 to 2011. At this time of initial contact, written consent to participate in the initial data collection, to be contacted for follow-up data collection, and longitudinal record review was provided by participants over the age of 18 or by the participant's parent or legal guardian for those under the age of 18. Written assent was also provided by participants under the age of 18. Participants provided written consent as adults to participate in additional data collection at follow-up timepoints. The University of New Mexico, Ethical and Independent Review Services/Salus IRB, the Office for Human Research Protections, and the juvenile correctional facility where data collection occurred (current IRB approval: Salus IRB; Study # 15050) approved of all research protocols. All research was performed in accordance with regulations set by the IRB.

In the second phase, outcomes data (i.e., both self-reported antisocial behavior post-release from the juvenile correctional facility and official adult charges and convictions) were collected. SWANC-Y study participants were followed-up with for up to 16 years after their release from the juvenile correctional facility to complete additional clinical assessments used to gauge involvement in antisocial behavior across the life course (see Assessments below; R01 HD092331, PI: Kiehl). Official arrests were extracted from criminal records obtained from the Administrative Office of the Courts in New Mexico and curated by the Center for Science and Law's Criminal Record Database (CRD⁴¹). Data in the CRD were matched with previous participants via four identifiers, including first and last name, date of birth, and social security number. In addition to these criminal records, online searches (social media, White Pages, Been Verified, county records, New Mexico Corrections Department offender search, and out of state inmate databases) were conducted to compile re-arrest data for all participants, including those not found in the CRD (i.e., participants who moved out of state).

Participants

The final sample for the present study included $n = 202$ incarcerated adolescent boys recruited from a maximum-security juvenile correctional facility in the state of New Mexico at the time of original data collection. Participants ranged from 13.82 to 19.37 years old ($M = 17.48$, $SD = 1.10$) at the time of their baseline MRI scan. Participants self-identified their race as American Indian or Alaskan Native ($n = 17$), Black or African American ($n = 11$), Native Hawaiian or Other Pacific Islander ($n = 1$), White ($n = 120$), or more than one race ($n = 6$), and their ethnicity as Hispanic or Latino ($n = 158$) or Non-Hispanic or Latino ($n = 36$). Additionally, $n = 47$ participants chose not to disclose their race, and $n = 8$ chose not to disclose their ethnicity. Participants reported medication use in the following categories at the time of MRI scan: antipsychotics ($n = 36$), antidepressants ($n = 79$), ADHD medications ($n = 32$), other ($n = 81$), unclassified ($n = 4$), and no medications ($n = 72$). Medications were considered unclassified if there was insufficient information via self-report or institutional file to determine the name and use. See Supplementary Table S1 for race and ethnicity counts, and Supplementary Table S2 for medication counts for participants included in both homicide and non-homicide groups.

Exclusion criteria for original data collection are as follows: traumatic brain injury associated with an extensive loss of consciousness, past or current history of CNS disease (e.g., stroke, multiple sclerosis, seizures, etc.), current or previous history of psychotic disorder as defined by the Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition (DSM-IV⁴²), first-degree relative with a history of psychotic disorder, hypertension or diabetes, mental retardation or fetal alcohol spectrum disorder, MRI contraindication (e.g., ferrous metal in body), low IQ (i.e., lower than 70), and low reading level (i.e., less than fourth-grade reading level). Participants were also excluded from the current study if they did not have a T1-weighted image. Finally, the current sample was restricted to male participants as there were too few female volunteers ($n = 57$) and the base-rate for future homicides among women was too low to power analyses ($n = 1$).

The total sample of $n = 202$ former SWANC-Y study participants was then further divided into two groups based on future homicidal behavior. The Homicide (H) group consisted of $n = 35$ participants who committed a homicide after their release from the juvenile correctional facility, whereas the No Homicide (No-H) group consisted of $n = 167$ participants who did not engage in homicidal behavior at any time point prior to incarceration as a youth or up to 16 years after their release from the juvenile correctional facility. This rate of homicide (i.e., 17%) is substantially higher than average rates of homicide committed by men reported in the state of New Mexico across a similar 16-year follow-up period (i.e., 0.0075%^{43,44}, follow-up period: 2007–

2023). Participants were included in the H group if they were charged with the following crimes after their release: Murder in any degree, Manslaughter, and/or Attempted Murder, with the exception of certain Vehicular Manslaughter convictions. In vehicular manslaughter cases, participants were excluded if the details of their crime indicated they did not willfully and intentionally kill another person (e.g., being involved in a car accident resulting in a death of another individual). Participants with charges of assault, battery, or other violent crimes remained in the No-H group because these charges do not indicate intent to cause immediate death or great bodily harm resulting in death. Charges were used to determine group assignment rather than convictions due to the high rate of charges being dropped or reduced in severity during plea bargaining in the United States. Using charges rather than final convictions therefore allowed us to determine the intent of violence based off of the official record of the crime. In addition to official charges, participants were also included in the H group if they self-reported engaging in homicide (i.e., committing homicide or attempted homicide in which an intent to kill was evident) via clinical interviews and self-report questionnaires obtained during follow-up data collection. The clinical interviews used to obtain self-reported homicide included the Hare Psychopathy Checklist: Youth Version (PCL: YV¹³), the Hare Psychopathy Checklist-Revised (PCL-R⁴⁵), and an in-house crime inventory (CI⁴⁶), collected at both the initial timepoint and at follow-up. Additionally, four assessments that were developed by our research group to capture antisocial behavior throughout the lifecourse were used to determine group membership. Specifically, two assessments were administered at the first timepoint as part of the original grant and the remaining two were longitudinal assessments administered during the follow-up period when participants were adults. A total of $n = 51$ participants (out of 202, 25%) completed at least one of these follow-up assessments. Information on self-reported homicidal behavior was cross-referenced between multiple assessments and timepoints to avoid incorrect group assignment due to false reporting. Based on these definitions, $n = 22$ participants had an official adult charge of homicide/manslaughter/attempted homicide and $n = 13$ participants self-reported engaging in homicidal behavior after their release from the juvenile correctional facility. Of the $n = 35$ participants in the H group, $n = 11$ previously committed homicide prior to their admission into the juvenile correctional facility. Within this grouping, $n = 3$ participants were classified in the homicide group in our previous report⁶. The remaining $n = 8$ were determined to have met current criteria for previous homicidal behavior based on additional information. Supplemental analyses were conducted excluding these $n = 11$ participants with previous homicides (see Supplementary Tables S9 and S10, and Fig. S3), to ensure reliability of results. Finally, participants were included in the No-H control group if they did not have an official charge or conviction of Murder in any degree, Manslaughter, or Attempted Murder in any degree, and did not self-report homicidal behavior at any timepoint (i.e., both before their admittance to the juvenile correctional facility and after their release).

Assessments

All baseline data included in analyses were collected from the SWANC-Y cohort as part of a larger study. The following variables were assessed: age at scan, IQ, socioeconomic status, number of substance use disorders, years of substance use, impulsivity, psychopathy, psychopathologies (i.e., anxiety, depression, attentional deficient-hyperactivity disorder, conduct disorder/oppositional defiant disorder, post-traumatic stress disorder), childhood trauma, number of traumatic brain injuries, age of first arrest, number of prior felonies, number of prior misdemeanors, number of prior probation/parole violations, total number of prior convictions, gang affiliation, parental incarceration, parental separation, total brain volume, and GMV estimates within regions of interest (ROIs; bilateral amygdala, insula, parahippocampal gyrus, middle temporal pole, superior temporal pole, orbitofrontal cortex). See Supplementary Information for additional details on all clinical variables.

MRI: Imaging Parameters

High-resolution T1-weighted structural MRI scans were acquired with the Mind Research Network Siemens 1.5T Avanto mobile MRI scanner, stationed at the juvenile correctional facility at the time of original data collection, using a multi-echo MPRAGE pulse sequence (repetition time = 2530 ms, echo times = 1.64 ms, 3.50 ms, 5.36 ms, 7.22 ms, inversion time = 1100 ms, flip angle = 7°, slice thickness = 1.3 mm, matrix size = 256 × 256) yielding 128 sagittal slices with an in-plane resolution of 1.0 × 1.0 mm.

MRI: Preprocessing and Region of Interest Selection

Data were pre-processed and analyzed using Statistical Parametric Mapping 12 (SPM12) software⁴⁷ (Wellcome Department of Cognitive Neurology, London, UK; <http://www.fil.ion.ucl.ac.uk/spm>). T1 images were automatically oriented to anterior-posterior commissure (AC-PC) alignment using the auto_acpc_reorient algorithm (https://github.com/lrq3000/auto_acpc_reorient) and were visually inspected to ensure proper realignment. Images were then spatially normalized to Montreal Neurological Institute (MNI) standard space via the Unified Segmentation approach as implemented in SPM12⁴⁸, which applies image registration based on Gaussian mixture modelling, tissue classification with warped prior probability maps, and bias correction to be combined in the same generative model. During spatial normalization, data were resampled to 2 × 2 × 2 mm, segmented into gray/white/cerebrospinal fluid maps, and modulated with Jacobian determinants to preserve total volume. Voxels with a gray matter value < 0.15 were excluded to remove possible partial volume effects. Finally, modulated and normalized segmented images were spatially smoothed with a 10 mm full width at half maximum (FWHM) Gaussian kernel.

Neuroanatomical regions of interest (ROIs) were selected based on previous literature documenting their association with psychopathic traits, antisocial behavior, and homicidal offending^{6,49–52}. These included regions including the bilateral amygdala, insula, parahippocampal gyrus, superior temporal pole, middle temporal pole, and orbitofrontal cortex. ROIs were delineated using the Wake Forest University (WFU) Pick Atlas Toolbox^{53,54}. These ROIs are displayed in Fig. 4.

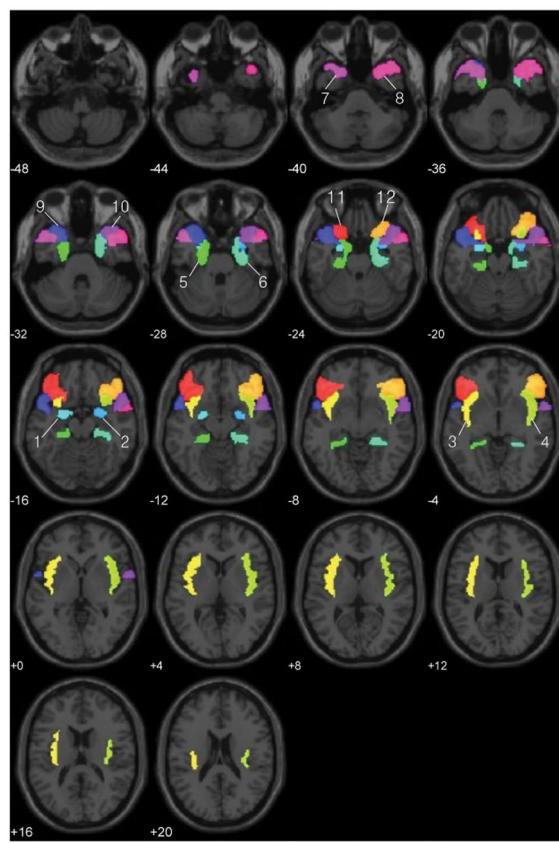


Fig. 4. The 12 a priori regions of interest. 1 and 2: left and right amygdala. 3 and 4: left and right insula. 5 and 6: left and right parahippocampal gyrus. 7 and 8: left and right middle temporal pole. 9 and 10: left and right superior temporal pole. 11 and 12: left and right orbitofrontal cortex. All regions were identified using the Wake Forest University (WFU) Pick Atlas Toolbox.

Data Analysis

Independent Samples *t*-test

Independent samples *t*-tests were conducted to investigate group differences between H and No-H participants on clinical and criminological variables. Due to a priori hypotheses regarding the relationships between homicide and psychopathic traits and age of first arrest, uncorrected one-tailed *t*-tests were performed for PCL: YV Total, Factor 1, and Factor 2 scores, and age of first arrest. Exploratory analyses compared additional variables (i.e., age at MRI scan, IQ, SES, Number of SUDs, ASI, BIS-11 total, number of TBIs, number of prior felonies, number of prior misdemeanors, number of prior parole/probation violations, total number of criminal convictions, total BV, and Trauma Checklist (TCL) Total, Factor 1, and Factor 2 scores) between groups, as these relationships are less supported. These post-hoc tests therefore underwent two-tailed independent samples *t*-tests, and a Bonferroni multiple comparison correction was utilized (i.e., 0.05/15, or $p < .003$).

Additional exploratory analyses were conducted comparing Future Homicide-Only (FH-Only; $n = 24$) boys, a subset of those in the H group who did not have a previous history of homicidal behavior prior to baseline data collection, to the No-H group on all aforementioned clinical variables. Further, independent samples *t*-tests were also conducted comparing the original H- and No-H participants on PCL: YV Facet and Item scores.

Fisher's Exact Test Fisher's Exact Tests were performed to investigate exploratory group differences between (a) H and No-H participants, and (b) FH-Only and No-H participants, on binary variables (i.e., all KSADS diagnoses, gang affiliation, parental incarceration, and parental separation). A Bonferroni multiple comparison correction was again utilized with these binary comparisons (i.e., 0.05/8, or $p = .006$).

Voxel-based morphometry

Voxel-based morphometry (VBM) is a voxel-wise neuroimaging technique, which was used here to determine group differences in gray matter volume (GMV) between H and No-H groups. An additional VBM analysis was conducted examining neuroanatomical differences between FH-Only and No-H groups. Specifically, two-sample *t*-tests were performed on a voxel-wise basis across gray matter segmentation maps. Given the sample sizes, we utilized non-parametric comparisons through the Randomise⁵⁵ algorithm in the FMRIB Software Library (FSL⁵⁶, applying $N = 10,000$ permutations, to generate contrast estimate maps. Brain volume (BV, sum of

gray matter and white matter volumes for each participant) and participant's age at MRI scan were included as covariates. Given the a priori selection of previously published ROIs, binarized masks of each hemisphere of the bilateral a priori ROIs (amygdala, insula, parahippocampal gyrus, superior temporal pole, middle temporal pole, and orbitofrontal cortex) were used for small volume correction (SVC) to spatially constrain the search area for peak estimates, using false discovery rate (FDR) correction at $p(\text{FDR}) \leq 0.05$ (Fig. 1). A final VBM analysis examining GVM differences between H- and No-H groups was conducted in which BV, age at MRI scan, and TCL scores were included as covariates.

Machine Learning

Weighted Linear Support Vector Machines (SVM), where the H and No-H groups are automatically weighted based on their sizes (inversely proportional to size in training data) was chosen to classify between the H and No-H groups due to the size imbalance between the two groups. This weighted linear SVM approach attempts to find the optimal decision boundary (hyperplane) that maximizes the margins between the H and No-H groups in the dataset feature space. The hyperparameter C in the SVM is a regularization term that balances the margin maximization and misclassification penalty. By implementing a weighted classifier, during the ML training procedure, the algorithms will choose the hyperparameter C of the SVM that assigns more importance (i.e., weight) to the correct identification of the minority homicide class members. This helps to reduce the classifier from being biased towards the majority non-homicide class. An a priori model included all ROIs, total BV, age at MRI scan, and PCL: YV factor scores. Additionally, the following models were conducted: (1) all clinical variables alone, (2) all neural data alone, and (3) all clinical and neural data combined for completeness (see Supplementary Table S11 for a list of all models investigated here). The value of the hyperparameter C was chosen using 1000 bootstrapped iterations randomly chosen from the range of 1e-5 to 1e-3 per training dataset ($n = 133$) and the best performing value on the validation dataset ($n = 33$) was chosen. This process was repeated 15 times across several splits of the 80% of the dataset ($n = 166$) that was reserved for model tuning and optimization. The most frequently chosen C value was subsequently used to train the 80% of the data and derive the best Linear SVM model. Homicide cases are intrinsically low probability events, hence, to maximize the identification of high-risk individuals over classification accuracy⁵⁷, instead of using resampling methods such as SMOTE, the default decision threshold of 0.5 for classification was altered. Additionally, since a significant number of the features are not well-separated [Supplementary Tables S5 and S6, Table 1], resampling methods may overfit to the minority class thereby reducing the generalizability of the classifier⁵⁸. Therefore, the decision threshold for determining the class membership in the best model was tuned to maximize the f1 score to balance precision and recall⁵⁹ to increase the classifiers prediction capability of the individuals in the Homicide class. The held-out test dataset ($n = 35$, H = 7, No-H = 28) was then applied to the tuned model to derive final performance. The tuned threshold classifier models generally improved the specificity of the models with imaging-variables with minimal loss of sensitivity, thereby improving overall accuracy. Feature importance weights for all the features were extracted to understand the contribution of each of them to the overall model performance.

Data availability

The data presented in this article are not readily available due to the potential for re-identification of participants in the present sensitive population (incarcerated juveniles). Interested parties should contact Dr. Kent Kiehl (kkiehl@unm.edu) for data used in this report, which may be shared under a signed data sharing agreement.

Received: 15 July 2025; Accepted: 12 December 2025

Published online: 24 December 2025

References

- Anderson, D. A. The aggregate cost of crime in the united States. *J. Law Econ.* **64**(4), 857–885 (2021).
- Miller, T. R., Cohen, M. A., Swedler, D. I., Ali, B. & Hendrie, D. V. Incidence and costs of personal and property crimes in the USA, 2017. *J. Benefit Cost Anal.* **12**(1), 24–54. <https://doi.org/10.1017/bca.2020.36> (2021).
- Cook, P. J., Ludwig, J. & Braga, A. A. Criminal records of homicide offenders. *JAMA* **294**(5), 598–601. <https://doi.org/10.1001/jam.a.294.5.598> (2005).
- Moffitt, T. E. Male antisocial behaviour in adolescence and beyond. *Nat. Hum. Behav.* **2**, 177–186 (2018).
- Hill-Smith, A. J., Hugo, P., Hughes, P., Fonagy, P. & Hartman, D. Adolescents murderers: abuse and adversity in childhood. *J. Adolesc.* **25**, 221–230 (2002).
- Cope, L. M. et al. Abnormal brain structure in youth who commit homicide. *Neuroimage Clin.* **4**, 800–8007. <https://doi.org/10.1016/j.nicl.2014.05.002> (2014).
- DeLisi, M., Piquero, A. R. & Cardwell, S. M. The unpredictability of murder: juvenile homicide in the pathways to desistance study. *YVJJ* **14**(1), 26–42 (2016).
- DiCataldo, F. & Everett, M. Distinguishing juvenile homicide from violent juvenile offending. *Int. J. Offender Ther. Comp. Criminol.* **52**, 158–174 (2008).
- Loeber, R. et al. The prediction of violence and homicide in young men. *J. Consult Clin. Psychol.* **73**(6), 1074–1088. <https://doi.org/10.1037/0022-006X.73.6.1074> (2005).
- Ahonen, L., Loeber, R. & Pardini, R. The prediction of young homicide and violent offenders. *Justice Q.* **33**(7), 1265–1291. <https://doi.org/10.1080/07418825.2015.1081263> (2016).
- Fox, B. & DeLisi, M. Psychopathic killers: A meta-analytic review of the psychopathy-homicide nexus. *Aggress. Violent Behav.* **44**, 67–79. <https://doi.org/10.1016/j.avb.2018.11.005> (2019).
- Piquero, A. R. & Chung, H. L. On the relationships between gender, early onset, and the seriousness of offending. *J. Crim Justice.* **29**(3), 189–206. [https://doi.org/10.1016/S0047-2352\(01\)00084-8](https://doi.org/10.1016/S0047-2352(01)00084-8) (2001).
- Forth, A. E., Kosson, D. S. & Hare, R. D. *The Hare Psychopathy Checklist: Youth Version* (Multi-Health Systems, 2003).
- Arlot, S. & Lerasle, M. Choice of V for V-fold cross-validation in least-squares density Estimation. *J. Mach. Learn. Res.* **17**(208), 1–50 (2016).
- Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai* **14**(2), 1137–1145 (1995).

16. Bischl, B., Mersmann, O., Trautmann, H. & Weihs, C. Resampling methods for meta-model validation with recommendations for evolutionary computation. *Evol. Comput.* **20**(2), 249–275. https://doi.org/10.1162/EVCO_a_00069 (2012).
17. He, H. & Garcia, E. A. Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* **21**(9), 1263–1284. <https://doi.org/10.1109/T-KDE.2008.239> (2009).
18. Allen, C. H. et al. The utility of expert-rated and self-report assessments of youth psychopathic traits for predicting felony recidivism among formerly incarcerated youth. *YVJJ* **22**(4), 231–250 <https://doi.org/10.1177/15412040241256574> (2024).
19. Allen, C. H. et al. Psychopathy scores predict recidivism in High-risk youth: A Five-year Follow-up study. *Res. Child. Adolesc. Psychopathol.* **52**, 1089–1103. <https://doi.org/10.1007/s10802-024-01169-x> (2024).
20. Corrado, R. R., Vincent, G. M., Hart, S. D. & Cohen, I. M. Predictive validity of the psychopathy checklist: youth version for general and violent recidivism. *Behav. Sci. Law.* **22**, 5–22. <https://doi.org/10.1002/bls.574> (2004).
21. Vincent, G. M., Odgers, C. L., McCormick, A. V. & Corrado, R. R. The PCL: YV and recidivism in male and female juveniles: A follow-up into young adulthood. *Int. J. Law Psychiatry.* **31**, 287–296. <https://doi.org/10.1016/j.jilp.2008.04.012> (2008).
22. Pardini, D. A., Raine, A., Erickson, K. & Loebel, R. Lower amygdala volume in men is associated with childhood aggression, early psychopathic traits, and future violence. *Biol. Psychiatry.* **75**(1), 73–80. <https://doi.org/10.1016/j.biopsych.2013.04.003> (2014).
23. DeLisi, M., Umphress, Z. R. & Vaugh, M. G. The criminology of the amygdala. *Crim Justice Behav.* **36**(11), 1241–1252. <https://doi.org/10.1177/0093854809343119> (2009).
24. Olson, I. R., Plotzker, A. & Ezzyat, Y. The enigmatic Temporal pole: A review of findings on social and emotional processing. *Brain* **130**, 1718–1731. <https://doi.org/10.1093/brain/awm052> (2007).
25. Kiehl, K. A. et al. Age of Gray matters: neuroprediction of recidivism. *Neuroimage: Clin.* **19**, 813–823. <https://doi.org/10.1016/j.nic.2018.05.036> (2018).
26. Johanson, M., Vaurio, O., Tiihonen, J. & Lähteenluoma, M. A systematic literature review of neuroimaging of psychopathic traits. *Front. Psychiatry* **10**, 1027 (2020). <https://doi.org/10.3389/fpsyg.2019.01027> (2020).
27. Lim, L., Radua, J. & Rubia, K. Gray matter abnormalities in childhood maltreatment: A Voxel-Wise Meta-Analysis. *Am. J. Psychiatry.* **171**(8), 854–863. <https://doi.org/10.1176/appi.ajp.2014.13101427> (2014).
28. Shold, J. N. et al. Psychometric properties of the trauma checklist 2.0 and its predictive utility of felony re-offending among high-risk juvenile offenders. *Child. Adolesc. Psychiatry Ment Health* **17**, 111 (2023). <https://doi.org/10.1186/s13034-023-00656-1> (2023).
29. Borum, R., Lodewijks, H., Bartel, P. A. & Forth, A. E. Structured assessment of violence risk in youth (SAVRY). in *Handbook of Violence Risk Assessment* 73–90 (Routledge, 2011).
30. Orbis Partners. *General Youth Assessment and Screening Instrument (YASI)* (2000). <https://orbispartners.com>.
31. de Vries Robb  , M., Geers, M. C. K., Stapel, M., Hilterman, E. L. B. & de Vogel, V. *SAPROF–Youth version: Guidelines for the assessment of protective factors for violence risk in juveniles* (ed Van Der Hoeven Kliniek) (2015).
32. Burghart, M. et al. The structured assessment of protective factors for violence risk (SAPROF): A meta-analysis of its predictive and incremental validity. *Psychol. Assess.* **35**(1), 56–67. <https://doi.org/10.1037/pas0001184> (2023).
33. Kleeven, A. T. H., De Vries Robb  , M., Mulder, E. A. & Popma, A. Risk assessment in juvenile and young adult offenders: predictive validity of the SAVRY and SAPROF-YV. *Assessment* **29**, 181–197. <https://doi.org/10.1177/107319112095974> (2022).
34. Koh, L. L., Day, A., Klettke, B., Daffern, M. & Chu, C. M. The predictive validity of youth violence risk assessment tools: a systematic review. *Psychol. Crime. Law.* **26**(8), 776–796. <https://doi.org/10.1080/1068316X.2020.1734200> (2020).
35. Scott, T., Brown, S. L. & Skilling, T. A. Predictive and convergent validity of the youth assessment and screening instrument in a sample of male and female Justice-Involved youth. *Crim Justice Behav.* **46**, 811–831. <https://doi.org/10.1177/0093854819842585s> (2019).
36. Caldwell, M. F. & Van Rybroek, G. J. Reducing violence in serious juvenile offenders using intensive treatment. *Int. J. Law Psychiatry.* **28**(6), 622–636. <https://doi.org/10.1016/j.jilp.2004.07.001> (2005).
37. Caldwell, M. F. Treatment-related changes in behavioral outcomes of psychopathy facets in adolescent offenders. *Law Hum. Behav.* **35**, 275–287. <https://doi.org/10.1007/s10979-010-9239-z> (2010).
38. Caldwell, M. F., Skeem, J., Salekin, R. & Van Rybroek, G. Treatment response of adolescent offenders with psychopathy features: A 2-year follow-up. *Crim Justice Behav.* **33**(5), 571–596. <https://doi.org/10.1177/0093854806288176> (2006).
39. Steiger, V. R. et al. Pattern of structural brain changes in social anxiety disorder after cognitive behavioral group therapy: A longitudinal multimodal MRI study. *Mol. Psychiatry.* **22**, 1164–1171 (2017).
40. Gonz  lez-Alemany, E. et al. Brain structural change associated with cognitive behavioral therapy in maltreated children. *Brain Res.* **1825**, 148702 (2024).
41. Ormachea, P. A., Haarsma, G., Davenport, S. & Eagleman, D. M. A new criminal records database for large-scale analysis of policy and behavior. *J. Sci. Law.* **1**(1), 1–7 (2015).
42. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR* (American Psychiatric Association, 2000).
43. Federal Bureau of Investigation., Crime Data Explorer. (2025). <https://cde.ucr.cjis.gov/LATEST/webapp/#/pages/explorer/crime/shr>
44. Census Bureau, U. S. Resident Population in New Mexico [NMPOP]. (2025). <https://fred.stlouisfed.org/series/NMPOP>
45. Hare, R. D. *The Psychopathy Checklist Revised*, 2nd edn (Multi-Health Systems, Toronto, 2003).
46. Aharoni, E. & Kiehl, K. A. Evading justice: quantifying criminal success in incarcerated psychopathic offenders. *Crim Justice Behav.* **40**(6), 629–645. <https://doi.org/10.1177/0093854812463565> (2013).
47. Ashburner, J. et al. SPM 12 (Wellcome Trust Centre for Neuroimaging, 2014).
48. Ashburner, J. & Friston, K. J. *Unified Segmentation Neuroimage* **26**(3), 839–851 (2005).
49. Fabian, J. M. Neuropsychological and neurological correlates in violent and homicidal offenders: A legal and neuroscience perspective. *Aggress. Violent Behav.* **15**(3), 209–223. <https://doi.org/10.1016/j.avb.2009.12.004> (2010).
50. Ermer, E., Cope, L. M., Calhoun, V. D., Nyalakanti, P. K. & Kiehl, K. A. Aberrant paralimbic Gray matter in incarcerated male adolescents with psychopathic traits. *J. Am. Acad. Child. Adolesc. Psychiatry.* **52**, 94–103. <https://doi.org/10.1016/j.jaac.2012.10.013> (2013).
51. Kiehl, K. A. A cognitive neuroscience perspective on psychopathy: evidence for paralimbic system dysfunction. *Psychiatry Res.* **142**, 107–128. <https://doi.org/10.1016/j.psychres.2005.09.013> (2006).
52. Sajous-Turner, A. et al. Aberrant brain Gray matter in murderers. *Brain Imaging Behav.* **14**, 2050–2061. <https://doi.org/10.1007/s11682-019-00155-y> (2020).
53. Maldjian, J. A., Laurienti, P. J., Kraft, R. A. & Burdette, J. H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* **19**, 1233–1239 (2003).
54. Maldjian, J. A., Laurienti, P. J. & Burdette, J. H. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage* **21**(1), 450–455 (2004).
55. Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. Permutation inference for the general linear model. *NeuroImage* **92**(C), 381–397. <https://doi.org/10.1016/j.neuroimage.2014.01.060> (2014).
56. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. *FSL NeuroImage* **62**(2), 782–790 <https://doi.org/10.1016/j.neuroimage.2011.09.015> (2012).
57. van den Goorbergh, R., van Smeden, M., Timmerman, D. & Van Calster, B. The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression. *J. Am. Med. Inf. Assoc.* **29**(9), 1525–1534. <https://doi.org/10.1093/jamia/ocac093> (2022).

58. Napierala, K. & Stefanowski, J. Types of minority class examples and their influence on learning classifiers from imbalanced data. *J. Intell. Inf. Syst.* **46**, 563–597. <https://doi.org/10.1007/s10844-015-0368-1> (2016).
59. Hancock, J., Johnson, J. M. & Khoshgoftaar, T. M. A comparative approach to threshold optimization for classifying imbalanced data. in *2022 IEEE 8th International Conference on Collaboration and Internet Computing (CIC)* 135–142 (2022).
60. Weschler, D. *Weschler Intelligence Scale for Children—Fourth Edition: Technical and Interpretive Manual* (Psychological Corporation, 2003).
61. Weschler, D. *WAIS-III: Weschler Adult Intelligence Scale* (Psychological Corporation, 1997).
62. Kaufman, J., Birmaher, B. & Brent, D. Schedule for affective disorders and schizophrenia for school-aged children: present and lifetime version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child. Adolesc. Psychiatry* **37**(7), 980–988 (1997).
63. McLellan, A. T. et al. The fifth edition of the addiction severity index. *J. Subst. Abuse Treat.* **9**(3), 199–213. [https://doi.org/10.1016/0740-5472\(92\)90062-S](https://doi.org/10.1016/0740-5472(92)90062-S) (1992).
64. Maurer, J. M., Steele, V. R., Vincent, G. M., Rao, V., Calhoun, V. D., Kiehl, K. A. Adolescent Psychopathic Traits Negatively Relate to Hemodynamic Activity within the Basal Ganglia during Error-Related Processing. *J. AbnorM. Child Psychol.* **47**(12), 1917–1929 <https://doi.org/10.1007/s10802-019-00560-3> (2019).
65. Maurer, J. M., Paul, S., Anderson, N. E., Nyalakanti, P. K., Kiehl, K. A. Youth with elevated psychopathic traits exhibit structural integrity deficits in the uncinate fasciculus. *NeuroImage: Clinical* **26**, 102236 <https://doi.org/10.1016/j.nicl.2020.102236> (2020).
66. Patton, J. H., Stanford, M. S. & Barratt, E. S. Factor structure of the Barratt impulsivity scale. *J. Clin. Psychol.* **51**(6), 768–774. [https://doi.org/10.1002/1097-4679\(199511\)51:6<768::AID-JCLP2270510607%3E3.0.CO;2-1](https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607%3E3.0.CO;2-1) (1995).
67. King, N. S., Crawford, S., Wenden, F. J., Moss, N. E. & Wade, D. T. The Rivermead post concussion symptoms questionnaire: A measure of symptoms commonly experienced after head injury and its reliability. *J. Neurol.* **242**, 587–592. <https://doi.org/10.1007/BF00868811> (1995).
68. Baglivio, M. T. The assessment of risk to recidivate among a juvenile offending population. *J. Crim Justice.* **37**(6), 596–607. <https://doi.org/10.1016/j.jcrimjus.2009.09.008> (2009).
69. Shader, M. United States Department of Justice. Risk factors for delinquency: An overview. Report NCJ No. 207540 (2003).
70. Baglivio, M. T., Wolff, K. T., Piquero, A. R. & Epps, N. The relationship between adverse childhood experiences (ACE) and juvenile offending trajectories in a juvenile offender sample. *J. Crim Justice.* **43**(3), 229–241. <https://doi.org/10.1016/j.jcrimjus.2015.04.012> (2015).
71. Darby, P. J., Allan, W. D., Kashani, J. H., Hartke, K. L. & Reid, J. C. Analysis of 112 juveniles who committed homicide: characteristics and a closer look at family abuse. *J. Fam Violence.* **13**, 365–375 (1998).

Author contributions

Conceptualization: SNR, ARG, JMM, KAK. Methodology: SNR, ARG, DS, JMM, KH, KAK. Software: ARG, DS. Funding acquisition: KAK. Project administration: KAK Writing—original draft: SNR, ARG, DS, JMM Writing—review & editing: SNR, ARG, DS, JMM, KH, CLH, KAK.

Funding

This research was supported by grants from the National Institute of Mental Health (R01-MH071896, PI: KAK) and National Institute of Child Health and Human Development (R01 HD092331, PI: KAK).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-32782-5>.

Correspondence and requests for materials should be addressed to K.A.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025