HHS Public Access

Author manuscript

Mol Psychiatry. Author manuscript; available in PMC 2021 November 12.

Published in final edited form as:

Mol Psychiatry. 2021 September ; 26(9): 4931–4943. doi:10.1038/s41380-020-0771-z.

Neuropsychosocial Markers of Binge Drinking in Young Adults

Joshua L. Gowin[1], **Peter Manza**[2], **Vijay A. Ramchandani**[3], **Nora D. Volkow**[2,4]

[1]University of Colorado Anschutz Medical Campus, Departments of Radiology and Psychiatry, Aurora, CO

[2]National Institute on Alcohol Abuse and Alcoholism, Laboratory of Neuroimaging, Bethesda, MD

[3]National Institute on Alcohol Abuse and Alcoholism, Section on Human Psychopharmacology, Bethesda, MD

[4]National Institute on Drug Abuse, Bethesda, MD

Abstract

Binge drinking is associated with disease and death, and developing tools to identify risky drinkers could mitigate its damage. Brain processes underlies risky drinking, so we examined whether neural and psychosocial markers could identify binge drinkers. Reward is the most widely studied neural process in addiction, but processes such as emotion, social cognition, and self-regulation are also involved. Here we examined whether neural processes apart from reward contribute to predicting risky drinking behaviors. From the Human Connectome Project, we identified 177 young adults who binged weekly and 309 non-bingers. We divided the sample into a training and a testing set and used machine-learning algorithms to classify participants based on psychosocial, neural or both (neuropsychosocial) data. We also developed separate models for each of seven fMRI tasks used in the study. An ensemble model developed in the training dataset was then applied to the testing dataset. Model performance was assessed by the area under the receiver operating characteristic curve (AUC) and differences between models were assessed using DeLong's test. The three models performed better than chance in the test sample with the neuropsychosocial (AUC = 0.86) and psychosocial (AUC = 0.84) performing better than the neural model (AUC = 0.64). Two fMRI-based models predicted binge drinking status better than chance, corresponding to the social and language tasks. Models developed with psychosocial and neural variables could contribute as diagnostic tools to help classify risky drinkers. Since social and language fMRI tasks performed best among the neural discriminators (including those from gambling and emotion tasks), it suggests the involvement of a broader range of brain processes than those traditionally associated with binge drinking in young adults.

Address of Correspondence: Joshua Gowin, PhD, Department of Radiology, Research Complex 2, 12700 E 19th Avenue, Room P15-1215, Aurora, CO 80045, joshua.gowin@ucdenver.edu, +1 303 724 3769.

Disclosures

All authors reported no biomedical financial interests or potential conflicts of interest.

Keywords

language; social; alcohol; random forest; support vector machine; elastic net; vulnerability; brain; MRI

Introduction

In the United States there are an estimated 88,000 deaths annually associated with alcohol[1] and in Canada alcohol accounts for more emergency room visits than heart attacks[2]. A consistent predictor of future alcohol use disorder (AUD) is early alcohol use[3, 4], and especially early-onset binge drinking[5, 6]. Further, young adults at greater risk for AUD are more likely to reach binge-level intoxication when alcohol is freely available[7, 8]. Identifying individuals at high risk for alcohol problems, such as regular binge drinkers, remains an important goal. Theory and empirical evidence suggest that altered brain function and structure might underlie alcohol use problems[9–12], indicating that neural markers may help identify risky drinkers. Despite theoretical support for brain-based markers of risky drinking, using them to develop robust predictive tools for substance use outcomes has been challenging[13, 14], likely due to the small effect size of most biological variables[15, 16] and the small sample sizes used in most neuroimaging studies. Further, many studies examine a small range of variables, which increases the likelihood that relevant variables are overlooked. It therefore remains unclear what aspects of behavior as well as neural function and structure best predict risky drinking, and how well these measures characterize risky drinking relative to well-known risk factors like age of first alcohol use[17]. An important step toward identifying predictive markers is to identify neural and psychosocial correlates of problematic alcohol use in current risky drinkers.

Numerous psychosocial risk factors of AUD have been identified, including family history of substance use disorders[18], personal history of psychopathology such as anxiety or antisocial behavior[18, 19] and personality traits such as impulsivity[20]. Among adolescents, a large prospective study reported that romantic and sexual experiences, early alcohol use, smoking cigarettes, and breaking rules were associated with binge drinking[21]. Poor emotion processing and dependence on other people to cope with stress also predict problematic substance use[22]. These psychosocial risk factors can serve as a benchmark with which to compare neural markers.

Several neuroimaging studies have compared differences in brain structure and function between risky drinkers and healthy adults. For example, relative to healthy men, alcohol-dependent men showed less differentiation between reward-associated and neutral cues in the ventral striatum[11]. Greater levels of activation in the ventral striatum and ventromedial prefrontal cortex during risk-taking tasks were shown to predict onset of binge drinking[23]. There is also evidence that amygdala activation during emotion processing may be a predictive marker of AUD, particularly in combination with ventral striatal activation during a reward task[24]; individuals with low response to reward but high response to threat had greater risk. This study examined two neurocognitive domains, but most fMRI studies examine only one, making it challenging to determine how domains compare at

differentiating healthy and risky drinkers. While imaging studies have examined neural activation during tasks examining risk-taking, reward, emotion, and alcohol cues[25], other neurocognitive domains remain largely unexplored.

Here we examine a large, high quality dataset with fMRI data from seven tasks, neuroanatomical structural data, and a wide range of psychosocial measures. We examined whether neural variables would contribute predictive accuracy beyond well-validated psychosocial variables. We also examined which fMRI tasks elicited brain activation patterns that best demarcate risky drinkers. The fMRI tasks spanned seven domains, including emotion processing, working memory, motor, social, gambling, relational, and language. Based on evidence that binge drinkers show differences in neural processing of reward and emotion[24], we hypothesized that the data from the gambling and emotion tasks would perform best. We further hypothesized that neural variables would improve model performance relative to psychosocial variables alone.

Material & Methods

Participants

We searched the Human Connectome Project[26] database and identified 177 young adults (range: 22–35 years old) who reported binge drinking at least once a week for the previous year (128 males, 72.3%). Binge drinking was defined as 4 or more drinks for females, or 5 or more drinks for males on a single day[27]. Three-hundred-thirty-six individuals reported no binge drinking in the past year, but of these, 27 met criteria for lifetime history of either alcohol abuse or dependence and were excluded, leaving 309 non-binge drinkers (101 males, 32.7%, see Table 1). The Human Connectome Project recruited twins, and approximately half of the selected sample for this study were monozygotic or dizygotic twins. Participants provided written informed consent and all procedures were approved by the Washington University Institutional Review Board (IRB # 201204036; Title: ‘Mapping the Human Connectome: Structure, Function and Heritability’)

Psychosocial Measures

A battery of questionnaires and behavioral tasks were collected, including psychiatric history, substance use history, personality, cognition, emotion and social function, and health and demographic measures (for complete details, visit https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release)[26]. A list of all 238 variables used in this study is available in the Supplemental Materials.

Psychiatric History: Self-reported function, including positive and negative affect, stress, anxiety, and social support, was assessed with the Achenbach Adult Self-Report (ASR) for Ages 18–59[28]. Participants were also assessed for depressive episodes, conduct problems, and depressive symptoms using the semi-structured assessment for the genetic of alcoholism (SSAGA)[29]. Psychiatric history assessed lifetime history of a diagnosis.

Substance Use History: Participants completed 7-day retrospective reports for alcohol and tobacco use. The SSAGA was used to assess age of first alcohol use, age of first

cannabis use, nicotine dependence, years smoked, difficulty quitting nicotine, lifetime illicit substance uses (i.e. cocaine, opioids, sedatives, stimulants), and lifetime cannabis uses.

Emotion Recognition and Well-being: Emotion recognition was assessed with the Penn Emotion Recognition Test[30, 31], where participants saw 40 faces and were asked whether the face's emotion was happy, sad, angry, scared, or if it expressed no feeling. Emotional state was assessed with the NIH Toolbox for the Assessment of Neurological and Behavioral Function (NIH Toolbox; www.nihtoolbox.org)[32] and included questionnaires for negative affect, positive affect, life satisfaction, meaning and purpose, friendship, social support, stress, and self-efficacy.

Personality: Participants completed the 60-item Neuroticism/Extroversion/Openness Five Factor Inventory (NEO-FFI)[33].

Cognitive Performance: For future planning, a delay discounting task was administered to determine how rapidly a person devalues money to be received in the future relative to money available immediately[31]. Values of $200 and $40,000 were assessed. For sustained attention, the Short Penn Continuous Performance Test was administered, where participants saw a series of arrangements of lines and were asked to press a button when the lines formed either a number or a letter[30, 31]. Verbal episodic memory was assessed using Form A of the Penn Word Memory Test, which presents participants with 20 words that they are asked to remember[30, 31]; participants subsequently see 40 words, including the initial 20, they are asked if they have seen the words previously. Fluid intelligence was assessed with Form A of an abbreviated version of the Raven's Progressive Matrices[31, 34]. Spatial orientation was assessed using the Variable Short Penn Line Orientation Test[30, 31] that showed participants two lines with different orientations and asked to rotate one line so that it became parallel with the other. The NIH Toolbox[32] was also used to assess episodic memory, cognitive flexibility, inhibition, vocabulary, processing speed, working memory, and reading.

Health and demographics: Health measurements included weight, height, body mass index, blood pressure, endocrine problems, and percentage of blood volume from red blood cells, and sleep quality as assessed by the Pittsburg Sleep Quality Index[35]. Demographics measures included age, race, ethnicity, handedness, income, employment status, education, and relationship status.

Family History of Psychiatric and Neurologic Disorders: Participants indicated if their mother or father had a history of psychosis, depression, bipolar disorder, anxiety, drug or alcohol problems, Alzheimer's disease, Parkinson's disease, or Tourette's syndrome. Each outcome was recorded as a binary variable (yes or no). If no history was present in the mother and father, binary variables indicated this for each parent.

fMRI tasks

The Human Connectome Project chose tasks to cover as many cognitive processes possible in a relatively short amount of time to maximize scientific value while minimizing participant burden[36]. Tasks were chosen that targeted well-characterized neural systems

and elicited reliable patterns of activation over time in individuals, and that had consistently detectable activation patterns in most participants[36]. Each participant in the dataset for this study had usable data for each of the seven tasks.

Working Memory: Participants saw blocks of trials that consisted of pictures of places, tools, faces and body parts (non-mutilated parts of bodies with no "nudity"). Within each run, the four different stimulus types were presented in separate blocks. Also, within each run, half of the blocks use a 2-back working memory task and half use a 0-back working memory task (as a working memory comparison). The contrast of interest was 2-back > 0-back.

Gambling[37]: Participants played a card guessing game to win or lose money. Participants were told that a mystery card could be from 1 to 9. They were asked to indicate if they thought the card was higher or lower than 5 by pressing one of two buttons. Participants then saw the card and either: 1) a green up arrow with "$1" for reward trials, 2) a red down arrow with "-$0.50" for loss trials; or 3) the number "5" and a double-headed arrow for neutral trials. Trials were predetermined to be reward, loss, or neutral trials, and the program generated the card's number after the participant pressed the button to be congruent with their choice. The contrast of interest was reward > punishment.

Motor[38, 39]: Participants saw cues that asked them to either tap their left or right fingers, or squeeze their left or right toes, or move their tongue to map motor areas. Each block of a movement type lasted 12 seconds (10 movements). The contrast of interest was movement > baseline.

Language[40]: The task consisted of blocks of a story task and a math task. During the story blocks, participants heard stories adapted from Aesop's fables, followed by a 2-alternative forced choice question that asks participants about the topic of the story. The contrast of interest was story > math.

Social Cognition (Theory of Mind): Participants saw short video clips of shapes (e.g. circles) that either interacted socially or moved randomly on the screen[41, 42]. After each clip, participants judged whether the objects interacted or moved randomly. An interaction was defined such that it appeared as if the shapes were considering each other's feelings and thoughts. The contrast of interest was social > random.

Relational Matching-to-Sample[43]: The stimuli were 6 different shapes (e.g. hexagon) filled with 1 of 6 different textures (e.g. polka dots). In the relational processing condition, participants saw a pair of objects on the top and bottom of the screen. Each pair differed on either shape or texture. Participants had to determine if shape or texture differed across the top pair and then decide if the bottom pair differed along the same dimension (e.g., if the top pair differs in shape, does the bottom pair also differ in shape?). In a control matching condition, participants saw two objects at the top of the screen and one at the bottom, and in the middle of the screen either the word "shape" or "texture". They had to decide whether the left or right of the top objects matched the bottom object on the specified dimension.

Subject responded by pressing one of two buttons. The contrast of interest was relational > match.

Emotion Processing[44]: Participants saw two faces presented on the top of the screen and a face on the bottom. They pressed a button to indicate if the bottom face expressed the same emotion as the left or right face on top of the screen. The faces had either angry, fearful, or neutral expressions. The contrast of interest was angry/fearful faces > neutral faces.

Imaging and preprocessing of fMRI data

Imaging data were acquired on a 3T Siemens "Connectome Skyra" scanner. For parameters, see reference[26] or supplemental material. Preprocessing was conducted as part of the HCP "minimally preprocessed" dataset with the FMRIB software library[45]; for details see reference[46] or supplemental material.

General Linear Model of fMRI task data

The design of the general linear model is described in detail elsewhere[36] and in the supplement. For each task, we parcellated each individual's contrast file by taking the average contrast value of all surface vertices within each of 360 cortical regions in an HCP atlas[47] and, for subcortical regions, the average value of all voxels within 19 regions in the Gordon atlas[48], for a total of 379 parcels.

Structural MRI data

Structural MRI data were processed using FreeSurfer software[49] and metrics of cortical thickness were included from each brain region in the Glasser atlas[47].

Group Differences in Neuropsychosocial Variables

To assess how the groups differed for the 238 psychosocial and 3016 neural variables included in the models, we used Cohen's *d* to estimate effect size. For the psychosocial variables, an effect size greater than 0.23 corresponded to a *p*-value less than 0.05 for a two-sample t-test when performing false-discovery rate correction[50]. For the neural variables, we controlled family-wise error to correct for multiple comparisons using nonparametric permutation testing with FSL's PALM software. This analysis was to provide insight into what features contribute to classification and what differences may exist between groups.

Machine Learning Analysis

The goal of this analysis was to determine how well binge drinkers could be identified using a neuropsychosocial dataset (all variables), a psychosocial dataset (only non-brain variables), and a neural dataset (only brain variables). We also wanted to determine which fMRI task best predicted binge drinking, so datasets from each task were used to generate separate models.

A schematic of the analytic approach is provided in Figure 1 and code is available in the Supplement. We imputed missing values using a proximity matrix implemented with the randomForest package in R software, version 3.6.2. The first step was to create a training set (75% of the sample) and a testing set (25% of the sample) of participants using the

caret package (topepo.github.io/caret/index.html;[51]) in R. This resulted in a training sample containing 133 binge drinkers and 252 non-binge drinkers. The testing sample containing 44 binge drinkers and 77 non-binge drinkers.

Second, we used the caretEnsemble package to classify individuals in the training sample as binge or non-binge drinkers. We tested three widely used algorithms designed for problems where the number of variables exceeds the number of participants: radial support vector machine[52], elastic net[53], and random forest[54]. We used 10-fold cross-validation and tested multiple parameters for each algorithm using a grid search for optimal performance. Ten-fold cross validation was repeated twenty times to reduce variability and increase reliability of models[55]. For elastic net, we optimized the parameters alpha and lambda. For random forest, we optimized the number of features to be sampled per node[56]. For radial support vector machine, we optimized the cost parameter[57]. Additional details on models are reported in Supplemental Materials.

Third, we generated an ensemble machine learning model by creating a general linear model from the three algorithms at their optimal performance level (for details, see Supplement). The linear model was generated with 10-fold cross validation repeated 20 times. We measured model performance by the area under the curve (AUC) for a receiver operating characteristic plot, thus considering both sensitivity and specificity. An importance score was calculated for each variable during this step (for details, see Supplement).

Fourth, we applied the ensemble model to the testing sample. The model assigned each participant in the testing class a probability that the participant was a binge drinker.

Fifth, we performed receiver operating characteristic analysis using the pROC package in R. The independent variable was the probability assigned to each participant. The dependent variable was a participant's status as either a binge or a non-binge drinker. The 95% confidence interval was estimated using the DeLong method[58], using bootstrapping of random subgroups of the total sample with 100 iterations. Since an area under the curve of 0.5 for ROC analysis represents chance performance, we defined significance as when the lower bound of the 95% confidence interval was greater than 0.5. Comparison of ROC curves was conducted using the DeLong method[58]. We calculated Brier scores using R package s2dverification. We also calculated overall accuracy (i.e. proportion of participants who were classified correctly), sensitivity, and specificity using the probability of greater than 0.5 from the ensemble prediction.

We also wanted to control for some noteworthy group discrepancies in proportion of males and differences in marijuana use, so we identified 117 binge drinkers without a history of cannabis dependence and fewer than 100 lifetime uses of cannabis. We identified 117 non binge drinkers who were matched to the 117 binge drinkers on age, sex, education, socioeconomic status, and body mass index (see Supplemental materials).

To address the potential bias that twins may introduce to classification schemes, we examined performance of the neuropsychosocial data among twins only (N=250) and among unrelated participants only (N=233), with the assumption that if both models performed similarly that it would indicate that the influence of twins is negligible. To examine

correlation among variables and its potential influence on model performance, we generated factor scores for several domains and generated a correlation plot (see Supplement).

Results

Group differences

The largest group differences were for age at first alcohol use, lifetime marijuana uses, and rule-breaking behavior ($d > 0.5$, $p < 0.05$); binge drinkers started drinking at a younger age, consumed more marijuana, and engaged in more rule-breaking (Figure 2). Binge drinkers were also more likely to be male (Table 1), reported higher availability of friends and greater extraversion ($d > 0.46$, $p < 0.05$). The neural variables with the largest effect sizes were widely distributed around the brain (Figure 2), but none of these were significant after correcting for multiple comparisons.

Models from the neuropsychosocial, psychosocial, and neural datasets

To test whether neural data could improve predictions of binge-drinking relative to psychosocial models, we constructed models for neural data, psychosocial data, and neuropsychosocial data (Table 2). The results for the three algorithms and the ensemble model in the training set are presented in the Supplement. In the test set, the neuropsychosocial model had the best performance (AUC = 0.86, 95%CI: 0.79, 0.93), with accuracy of 0.80, sensitivity of 0.70, and specificity of 0.86. The psychosocial model (AUC = 0.84, 95%CI: 0.76, 0.91) performed similarly (z= 0.76, $p = 0.446$), with accuracy of 0.79, sensitivity of 0.80, and specificity of 0.79. Both the neuropsychosocial (z = 4.42, $p < 0.001$) and the psychosocial (z = 3.08, $p = 0.002$) models performed significantly better than the neural model (Figure 3), but the neural model was significantly better than chance (AUC = 0.64, 95%CI: 0.54, 0.74). The neural model had accuracy of 0.64 and specificity of 0.83, but sensitivity of 0.30, suggesting that the decline in accuracy was largely due to a decrease in sensitivity, with specificity largely unchanged. To control for group differences in age and proportion of males, we performed an analysis on subgroups of binge and non-binge drinkers, but found similar AUC (see Supplemental Materials), suggesting that the model performance we achieved was not simply due to group differences.

To examine the influence of substance use history, the neuropsychosocial and psychosocial models were examined without these variables. The difference in model performance was not significant ($p > 0.05$, see Table 2 and Figure 3). To examine the influence of relatedness of participants, we examined all twins and all unrelated participants separately. The AUC of the ensemble model for twins (0.88) did not differ ($p > 0.05$) from the model for unrelated participants (0.89), suggesting that including related participants did not substantially bias our models. Most domains of the neuropsychosocial model had relatively low correlations, suggesting that collinearity between fMRI variables does not account for the lack of improvement of the model of the neuropsychosocial model over the psychosocial model (see Supplement).

Models from the fMRI tasks

We tested the performance of data from seven fMRI tasks in the training sample. Four tasks produced models that performed no better than chance (the lower bound of the 95% confidence interval included 0.5), including relational matching-to-sample, gambling, emotion, and working memory tasks; these tasks were not considered further. The language, motor, and social tasks performed above chance in the training data, so ensemble models from these datasets were applied to the test set. The motor task model did not perform significantly better than chance (AUC = 0.58, 95% CI 0.48, 0.69) but the language (AUC = 0.70, 95% CI: 0.60, 0.80) and social (AUC = 0.62, 95% CI: 0.52, 0.73) models outperformed chance. Delong's test revealed no significant difference in the performance of the four models (Figure 4; $p > 0.05$). Accuracy of the three models was as high as 0.66 for the language task and as low as 0.63 for the motor task (Table 2), with all models having sensitivity below 0.33 but having specificity above 0.85.

Variable importance

Variable importance for the neuropsychosocial model indicated that cortical thickness in the parietal and frontal cortex contributed to model performance, in addition to marijuana use and relationship status (see Supplemental materials). The psychosocial model relied on smoking variables and rule breaking behavior. When substance history was excluded, friendship availability became one of the most important predictors.

Variable importance scores were calculated for the social, and language tasks (Figure 5 and Supplemental Materials). The most important regions for the language task included the left premotor cortex and the left anterior cingulate cortex. For the social task, the most important regions included the right pars opercularis and the left lateral occipital cortex. For each task, numerous brain regions contributed to classification accuracy, suggesting the need for models that consider regions from multiple brain networks.

Discussion

This study examined whether binge drinking status in young adults could be identified using neuropsychosocial data. The neural model performed better than chance, but contrary to our hypothesis the neural variables did not lead to significant improvements in classification relative to the psychosocial variables, which likely reflects the strong predictive value of the psychosocial variables. In particular, early age of first alcohol use has been noted by others as a strong predictor of binge drinking[21]. The results did not support our hypothesis that the gambling and emotion fMRI tasks would identify binge drinkers. Instead, we found that only the language and social tasks performed significantly better than chance, suggesting the relevance of neuronal networks related to language and social processing in the risk for AUD.

Machine learning performance was generally successful, and we replicated that the elastic net model produced the strongest results in predicting alcohol use, relative to SVM and random forest[59]. History variables, most prominently age at first drink, and rule-breaking behavior, showed the largest difference between the binge and non-binge drinkers in this

sample, and contributed most to successful classification. These differences in young adults parallel the importance of history in classifying binge drinking adolescents[21]. While this study established variables that are correlates of risky drinking, it is likely that some of the variables precede the onset of risky drinking. Causal risk factors are variable markers that, if changed, alter the risk of developing risky drinking. Although age of first drink is a marker that precedes risky drinking, twin studies have suggested that delaying the age of first drink would not prevent risky drinking, and thus it is not a causal risk factor[17]. Biological variables that are present prior to onset of heavy drinking, like genes and brain function, are more likely to be causal risk factors, such that altering them could decrease the likelihood of developing risky drinking patterns[60, 61]. Identifying neural functioning differences between binge drinkers and non-binge drinkers could offer novel targets for therapy and reduction of risky drinking.

Two of the seven fMRI tasks produced significant models, corresponding to the social and language tasks. Similarly the psychosocial data offered support for group differences in language and social function, as binge drinkers had poorer verbal episodic memory than non-binge drinkers. Previous studies have also found language-related impairments associated with problematic substance use[62]. The differences in brain activity during the language task may relate to poorer performance on language tests[63]. Similarly, differences in neural processing of social interactions between groups was corroborated by differences in self-reported friendship availability and extraversion. Social and language processing differences suggest that linguistic and socio-emotional factors could contribute to drinking behavior; for example, individuals from low relative to high socioeconomic backgrounds are more likely to binge drink[64] and are also more likely to prefer communicating through text message instead of voice calls[65]. Text preference may have clinical relevance, as text-message interventions that use realistic language are effective in reducing binge drinking among the socially disadvantaged[66]. Future studies could implement methods that identify common sources of variation linking brain function with behavior (e.g., canonical correlation analysis) to provide a more comprehensive picture on how language and social processing are associated with risky drinking behavior. Though none of the neural variables for the fMRI task reached significance in distinguishing between groups, the five regions identified in the language fMRI task were lateralized and localized in frontal (*e.g.* anterior cingulate cortex) and parietal regions and the two regions for the social fMRI task were located in right superior and inferior temporal cortices. Some of these brain regions differ between controls and alcoholics[67]. However, further work is needed to understand the neural correlates of social and language processing since they could serve as potential targets of interventions aimed at reducing risky drinking.

We failed to find an association between some traits with a well-established link to risky drinking. For example, a number of well-powered studies have found that low levels of conscientiousness is associated with higher levels of risky drinking in both adolescents[68] and adults[69, 70], but we did not observe that association in this study. Given the strength and consistency of this relationship in other studies, we were surprised by its absence here, which may reflect idiosyncrasies of the sample. Delay discounting, an index of choice impulsivity, has also been associated with risky alcohol use[71], but did not differ significantly between groups in the present study.

There are several limitations to this study. The study was cross-sectional, so it is unclear if any differences between binge and non-binge drinkers resulted from drinking or were present prior to drinking. A large proportion of the participants had a twin in the study, potentially biasing our results, but our evidence suggests that any potential bias was minimal. This study used a definition of binge drinking that did not include diagnosis of AUD, so these results may not generalize to individuals who have alcohol problems due to high levels of craving or negative consequences due to drinking. Further, it is possible that some non-binge drinker participants may have had periods of weekly binge drinking earlier in their lifetime, which would decrease the sensitivity to distinguish them from current binge drinkers. However, this is unlikely, since we excluded individuals with a lifetime history of alcohol abuse or dependence in our non-binge drinker sample (Table 1). Since there is likely overlap between weekly binge drinking and meeting criteria for abuse, excluding those with a lifetime history of AUD would also remove many with a history of weekly binge drinking. Additionally, the high model performance suggests that if there were prior binge drinkers in the comparison group, their presence did not prevent good classification. The imaging tasks were not chosen to assess substance-specific neural functions. Future studies that use alcohol-specific fMRI tasks, such as alcohol-cue reactivity paradigms, may find stronger evidence of neural differences between binge and non-binge drinkers. Despite the relatively large sample size, binge drinkers have substantial heterogeneity, so improving the accuracy of prediction models based on neuroimaging variables may require larger datasets. This sample was mostly male, so our models may be less applicable to females. However, we did examine all the results separately for males and females and did not observe any noteworthy differences in model performance across sex, but further testing is required to confirm this. Further, the racial and ethnic makeup of the two groups differed, and since we did not control for this in the analysis, it is possible that factors associated with race and ethnicity affected our results.

This study shows that psychosocial variables can classify risky drinkers with high accuracy, with an AUC approaching 0.9, which approximates the levels of a diagnostic test. While neural variables did not approach this high level of accuracy, they produced models that outperformed chance, suggesting that neuroimaging may be developed as a tool to characterize AUD diagnosis or to help clinicians make decisions about treatment courses for patients. Due to cost, MRI (or Positron Emission Tomography) are unlikely to become common in typical clinical settings, but the possibility of them providing clinically relevant information could nonetheless influence AUD treatment. By elucidating the neural basis of AUD, including the complexity and heterogeneity of the disorder, we may provide hope for individuals unresponsive to standard treatment, or we may identify new therapeutic targets.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 National Institutes of Health (NIH) institutes and centers that support the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems

Neuroscience at Washington University. We thank Francisco Pereira for suggestions on setting up the analysis. Support for time on this manuscript was provided by the National Institute on Alcohol Abuse and Alcoholism (Z1A AA000466 to VAR and R00 AA024778 to JLG).

References

1. Alcohol and Public Health: Alcohol-Related Disease Impact (ARDI). Average for United States 2006–2010 Alcohol-Attributable Deaths Due to Excessive Alcohol Use. https://www.cdc.gov/alcohol/onlinetools.htm, 2018, Accessed Date Accessed 2018 Accessed.
2. Alcohol Harm in Canada Examining Hospitalizations Entirely Caused by Alcohol and Strategies to Reduce Alcohol Harm. https://www.cihi.ca/sites/default/files/document/report-alcohol-hospitalizations-en-web.pdf, 2017, Accessed Date Accessed 2017 Accessed.
3. Hingson RW, Heeren T, Winter MR. Age at drinking onset and alcohol dependence: age at onset, duration, and severity. Archives of pediatrics & adolescent medicine 2006; 160(7): 739–746. [PubMed: 16818840]
4. Dawson DA, Goldstein RB, Chou SP, Ruan WJ, Grant BF. Age at first drink and the first incidence of adult-onset DSM-IV alcohol use disorders. Alcoholism, clinical and experimental research 2008; 32(12): 2149–2160.
5. Kuntsche E, Rossow I, Simons-Morton B, Bogt TT, Kokkevi A, Godeau E. Not early drinking but early drunkenness is a risk factor for problem behaviors among adolescents from 38 European and North American countries. Alcoholism, clinical and experimental research 2013; 37(2): 308–314.
6. Chassin L, Pitts SC, Prost J. Binge drinking trajectories from adolescence to emerging adulthood in a high-risk sample: predictors and substance abuse outcomes. J Consult Clin Psychol 2002; 70(1): 67. [PubMed: 11860058]
7. Gowin JL, Sloan ME, Stangl BL, Vatsalya V, Ramchandani VA. Vulnerability for Alcohol Use Disorder and Rate of Alcohol Consumption. The American journal of psychiatry 2017; 174(11): 1094–1101. [PubMed: 28774194]
8. Sloan ME, Gowin JL, Janakiraman R, Ester CD, Stoddard J, Stangl B et al. High-risk social drinkers and heavy drinkers display similar rates of alcohol consumption. Addict Biol 2019.
9. Koob GF, Volkow ND. Neurocircuitry of addiction. Neuropsychopharmacology 2010; 35(1): 217. [PubMed: 19710631]
10. Makris N, Oscar-Berman M, Jaffin SK, Hodge SM, Kennedy DN, Caviness VS et al. Decreased Volume of the Brain Reward System in Alcoholism. Biological Psychiatry 2008; 64(3): 192–202. [PubMed: 18374900]
11. Beck A, Schlagenhauf F, Wüstenberg T, Hein J, Kienast T, Kahnt T et al. Ventral Striatal Activation During Reward Anticipation Correlates with Impulsivity in Alcoholics. Biological Psychiatry 2009; 66(8): 734–742. [PubMed: 19560123]
12. Bjork JM, Smith AR, Hommer DW. Striatal sensitivity to reward deliveries and omissions in substance dependent patients. NeuroImage 2008; 42(4): 1609–1621. [PubMed: 18672069]
13. Gowin JL, Ball TM, Wittmann M, Tapert SF, Paulus MP. Individualized relapse prediction: Personality measures and striatal and insular activity during reward-processing robustly predict relapse. Drug and Alcohol Dependence 2015; 152: 93–101. [PubMed: 25977206]
14. Gowin JL, Ernst M, Ball T, May AC, Sloan ME, Tapert SF et al. Using neuroimaging to predict relapse in stimulant dependence: A comparison of linear and machine learning models. NeuroImage: Clinical 2019; 21: 101676. [PubMed: 30665102]
15. Kwako LE, Schwandt ML, Ramchandani VA, Diazgranados N, Koob GF, Volkow ND et al. Neurofunctional Domains Derived From Deep Behavioral Phenotyping in Alcohol Use Disorder. The American journal of psychiatry 2019: appi. ajp. 2018.18030357.
16. Sloan ME, Gowin JL, Yan J, Schwandt ML, Spagnolo PA, Sun H et al. Severity of alcohol dependence is associated with the fatty acid amide hydrolase Pro129Thr missense variant. Addict Biol 2018; 23(1): 474–484. [PubMed: 28150397]
17. Prescott CA, Kendler KS. Age at First Drink and Risk for Alcoholism: A Noncausal Association. Alcoholism: Clinical and Experimental Research 1999; 23(1): 101–107.

18. Cheng ATA, Gau S-F, Chen THH, Chang J-C, Chang Y-T. A 4-Year Longitudinal Study on Risk Factors for Alcoholism. JAMA Psychiatry 2004; 61(2): 184–191.
19. Poikolainen K RISK FACTORS FOR ALCOHOL DEPENDENCE: A CASE-CONTROL STUDY. Alcohol and Alcoholism 2000; 35(2): 190–196. [PubMed: 10787396]
20. Granö N, Virtanen M, Vahtera J, Elovainio M, Kivimäki M. Impulsivity as a predictor of smoking and alcohol consumption. Personality and Individual Differences 2004; 37(8): 1693–1700.
21. Whelan R, Watts R, Orr CA, Althoff RR, Artiges E, Banaschewski T et al. Neuropsychosocial profiles of current and future adolescent alcohol misusers. Nature 2014; 512(7513): 185. [PubMed: 25043041]
22. Wong CF, Silva K, Kecojevic A, Schrager SM, Bloom JJ, Iverson E et al. Coping and emotion regulation profiles as predictors of nonmedical prescription drug and illicit drug use among high-risk young adults. Drug and Alcohol Dependence 2013; 132(1): 165–171. [PubMed: 23453258]
23. Morales AM, Jones SA, Ehlers A, Lavine JB, Nagel BJ. Ventral striatal response during decision making involving risk and reward is associated with future binge drinking in adolescents. Neuropsychopharmacology 2018; 43(9): 1884–1890. [PubMed: 29789576]
24. Nikolova YS, Knodt AR, Radtke SR, Hariri AR. Divergent responses of the amygdala and ventral striatum predict stress-related problem drinking in young adults: possible differential markers of affective and impulsive pathways of risk for alcohol use disorder. Mol Psychiatry 2016; 21(3): 348. [PubMed: 26122584]
25. Courtney AL, Rapuano KM, Sargent JD, Heatherton TF, Kelley WM. Reward System Activation in Response to Alcohol Advertisements Predicts College Drinking. Journal of studies on alcohol and drugs 2018; 79(1): 29–38. [PubMed: 29227227]
26. Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TE, Bucholz R et al. The Human Connectome Project: a data acquisition perspective. Neuroimage 2012; 62(4): 2222–2231. [PubMed: 22366334]
27. Wechsler H, Davenport A, Dowdall G, Moeykens B, Castillo S. Health and behavioral consequences of binge drinking in college: A national survey of students at 140 campuses. JAMA 1994; 272(21): 1672–1677. [PubMed: 7966895]
28. Achenbach TM. Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications. University of Vermont, Research Center of Children, Youth & Families2009.
29. Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr. et al. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. J Stud Alcohol 1994; 55(2): 149–158. [PubMed: 8189735]
30. Gur RC, Ragland JD, Moberg PJ, Turner TH, Bilker WB, Kohler C et al. Computerized Neurocognitive Scanning:: I. Methodology and Validation in Healthy People. Neuropsychopharmacology 2001; 25(5): 766–776. [PubMed: 11682260]
31. Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB et al. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. Journal of neuroscience methods 2010; 187(2): 254–262. [PubMed: 19945485]
32. Gershon RC, Cella D, Fox NA, Havlik RJ, Hendrie HC, Wagster MV. Assessment of neurological and behavioural function: the NIH Toolbox. The Lancet Neurology 2010; 9(2): 138–139. [PubMed: 20129161]
33. McCrae RR, Costa PT Jr. A contemplated revision of the NEO Five-Factor Inventory. Pers Individ Dif 2004; 36(3): 587–596.
34. Bilker WB, Hansen JA, Brensinger CM, Richard J, Gur RE, Gur RC. Development of abbreviated nine-item forms of the Raven's standard progressive matrices test. Assessment 2012; 19(3): 354–369. [PubMed: 22605785]
35. Buysse DJ, Reynolds III CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989; 28(2): 193–213. [PubMed: 2748771]

36. Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M et al. Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage 2013; 80: 169–189. [PubMed: 23684877]
37. Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA. Tracking the hemodynamic responses to reward and punishment in the striatum. Journal of neurophysiology 2000; 84(6): 3072–3077. [PubMed: 11110834]
38. Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BT. The organization of the human cerebellum estimated by intrinsic functional connectivity. J Neurophysiol 2011; 106(5): 2322–2345. [PubMed: 21795627]
39. Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 2011; 106(3): 1125–1165. [PubMed: 21653723]
40. Binder JR, Gross WL, Allendorfer JB, Bonilha L, Chapin J, Edwards JC et al. Mapping anterior temporal lobe language areas with fMRI: a multicenter normative study. NeuroImage 2011; 54(2): 1465–1475. [PubMed: 20884358]
41. Castelli F, Happé F, Frith U, Frith C. Movement and Mind: A Functional Imaging Study of Perception and Interpretation of Complex Intentional Movement Patterns. NeuroImage 2000; 12(3): 314–325. [PubMed: 10944414]
42. Wheatley T, Milleville SC, Martin A. Understanding Animate Agents: Distinct Roles for the Social Network and Mirror System. Psychological Science 2007; 18(6): 469–474. [PubMed: 17576256]
43. Smith R, Keramatian K, Christoff K. Localizing the rostrolateral prefrontal cortex at the individual level. NeuroImage 2007; 36(4): 1387–1396. [PubMed: 17532648]
44. Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, Goldman D et al. Serotonin transporter genetic variation and the response of the human amygdala. Science 2002; 297(5580): 400–403. [PubMed: 12130784]
45. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. NeuroImage 2012; 62(2): 782–790. [PubMed: 21979382]
46. Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL et al. The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage 2013; 80: 105–124.
47. Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E et al. A multi-modal parcellation of human cerebral cortex. Nature 2016; 536(7615): 171–178. [PubMed: 27437579]
48. Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cerebral cortex (New York, NY : 1991) 2016; 26(1): 288–303.
49. Fischl B FreeSurfer. NeuroImage 2012; 62(2): 774–781. [PubMed: 22248573]
50. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. 1995; 57(1): 289–300.
51. Kuhn M Building predictive models in R using the caret package. J Stat Softw 2008; 28(5): 1–26. [PubMed: 27774042]
52. Scholkopf B, Sung K-K, Burges CJ, Girosi F, Niyogi P, Poggio T et al. Comparing support vector machines with Gaussian kernels to radial basis function classifiers. IEEE Trans Signal Process 1997; 45(11): 2758–2765.
53. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodol 2005; 67(2): 301–320.
54. Breiman L Random forests. Machine Learning 2001; 45(1): 5–32.
55. Varoquaux G Cross-validation failure: Small sample sizes lead to large error bars. NeuroImage 2018; 180: 68–77. [PubMed: 28655633]
56. Goldstein BA, Hubbard AE, Cutler A, Barcellos LF. An application of Random Forests to a genome-wide association dataset: methodological considerations & new findings. BMC Genetics 2010; 11(1): 49. [PubMed: 20546594]
57. Efficient parameter selection for support vector machines in classification and regression via model-based global optimization. Proceedings of the Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005. 2005. IEEE.

58. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988: 837–845. [PubMed: 3203132]
59. Afzali MH, Sunderland M, Stewart S, Masse B, Seguin J, Newton N et al. Machine-learning prediction of adolescent alcohol use: a cross-study, cross-cultural validation. Addiction (Abingdon, England) 2019; 114(4): 662–671.
60. Kirsch M, Gruber I, Ruf M, Kiefer F, Kirsch P. Real-time functional magnetic resonance imaging neurofeedback can reduce striatal cue-reactivity to alcohol stimuli. Addict Biol 2016; 21(4): 982–992. [PubMed: 26096546]
61. Yedy I, Quintanilla ME, Ezquer F, Morales P, Rivera-Meza M, Karahanian E et al. Gene and cell therapy on the acquisition and relapse-like binge drinking in a model of alcoholism: translational options. Gene therapy 2019; 26(10–11): 407–417. [PubMed: 30820030]
62. Tapert SF, Brown SA. Substance dependence, family history of alcohol dependence and neuropsychological functioning in adolescence. Addiction (Abingdon, England) 2000; 95(7): 1043–1053.
63. Buchweitz A, Mason RA, Tomitch L, Just MA. Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. Psychol Neurosci 2009; 2(2): 111–123. [PubMed: 21526132]
64. Fone DL, Farewell DM, White J, Lyons RA, Dunstan FD. Socioeconomic patterning of excess alcohol consumption and binge drinking: a cross-sectional study of multilevel associations with neighbourhood deprivation. BMJ open 2013; 3(4).
65. Smith A Americans and text messaging, vol. 24. Pew Internet & American Life Project Washington, DC: Pew Research Center, 2011.
66. Irvine L, Melson AJ, Williams B, Sniehotta FF, McKenzie A, Jones C et al. Real Time Monitoring of Engagement with a Text Message Intervention to Reduce Binge Drinking Among Men Living in Socially Disadvantaged Areas of Scotland. International journal of behavioral medicine 2017; 24(5): 713–721. [PubMed: 28702758]
67. Mackey S, Allgaier N, Chaarani B, Spechler P, Orr C, Bunn J et al. Mega-Analysis of Gray Matter Volume in Substance Dependence: General and Substance-Specific Regional Effects. The American journal of psychiatry 2019; 176(2): 119–128. [PubMed: 30336705]
68. Hagger-Johnson G, Bewick BM, Conner M, O'Connor D, Shickle D. School-related conscientiousness, alcohol drinking, and cigarette smoking in a representative sample of English school pupils. British journal of health psychology 2012; 17(3): 644–665. [PubMed: 22353559]
69. Bogg T, Roberts BW. Conscientiousness and health-related behaviors: a meta-analysis of the leading behavioral contributors to mortality. Psychological bulletin 2004; 130(6): 887–919. [PubMed: 15535742]
70. Cook M, Young A, Taylor D, Bedford AP. Personality correlates of alcohol consumption. Personality and Individual Differences 1998; 24(5): 641–647.
71. MacKillop J, Amlung MT, Few LR, Ray LA, Sweet LH, Munafo MR. Delayed reward discounting and addictive behavior: a meta-analysis. Psychopharmacology 2011; 216(3): 305–321. [PubMed: 21373791]

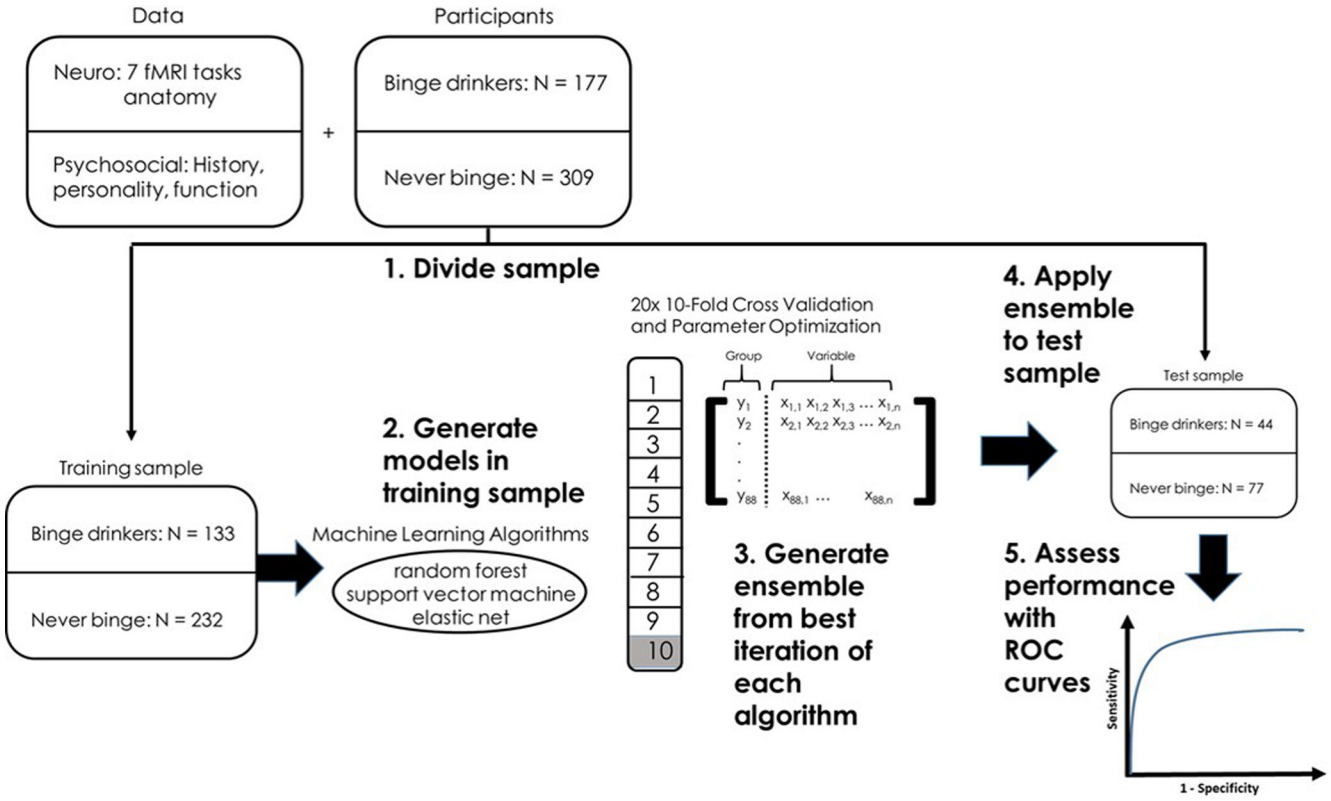

Figure 1. Schematic of our analytic approach.

Data were comprised of fMRI data from seven tasks, neuroanatomical data, and a range of data assessing personality, cognition, history, and more. The sample was divided into a training and test sample. Models were generated in the training sample using random forest, radial support vector machine, and elastic net algorithms. Tenfold cross-validation was repeated 20 times and was optimized on model parameters to maximize area under the receiver operating characteristic curve. The best performing model from each algorithm was used to generate an ensemble model using general linear regression, and this was also repeated with tenfold cross-validation. The ensemble was applied to the test sample. The performance in the test sample was assessed by the area under the receiver operating characteristic curve.

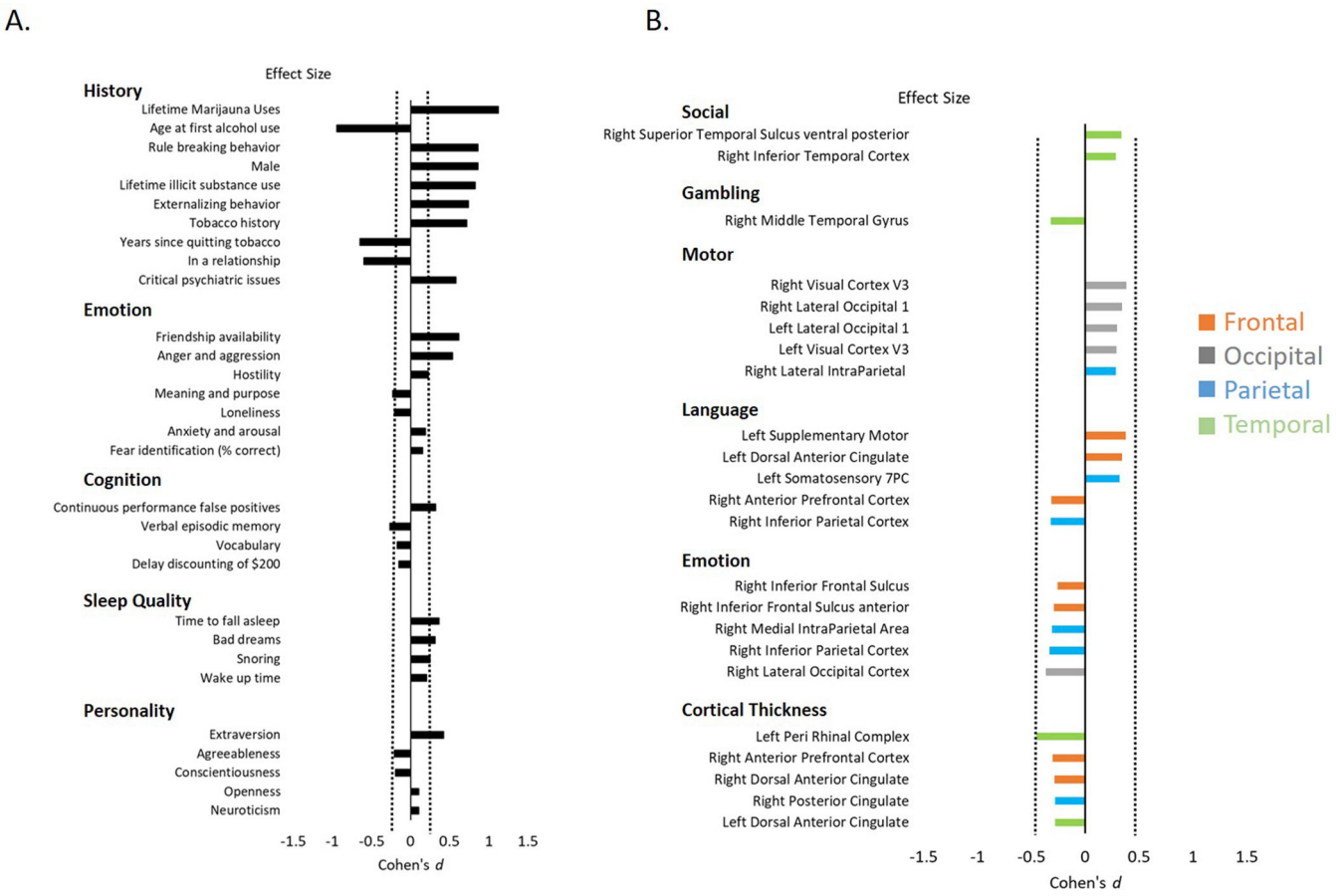

Figure 2. Group differences.

Effect sizes (Cohen's *d*) for the two sample *t*-tests comparing binge drinkers (n=177) with non-binge drinkers (n=309). Panel A presents psychosocial variables and panel B presents neural variables. Panel B is color-coded to highlight the brain lobes where differences exist. Positive values for *d* indicate significantly higher values in the binge drinking group relative to controls. The dashed line represent *p*-value < 0.05 using false discovery rate correction for multiple comparisons.

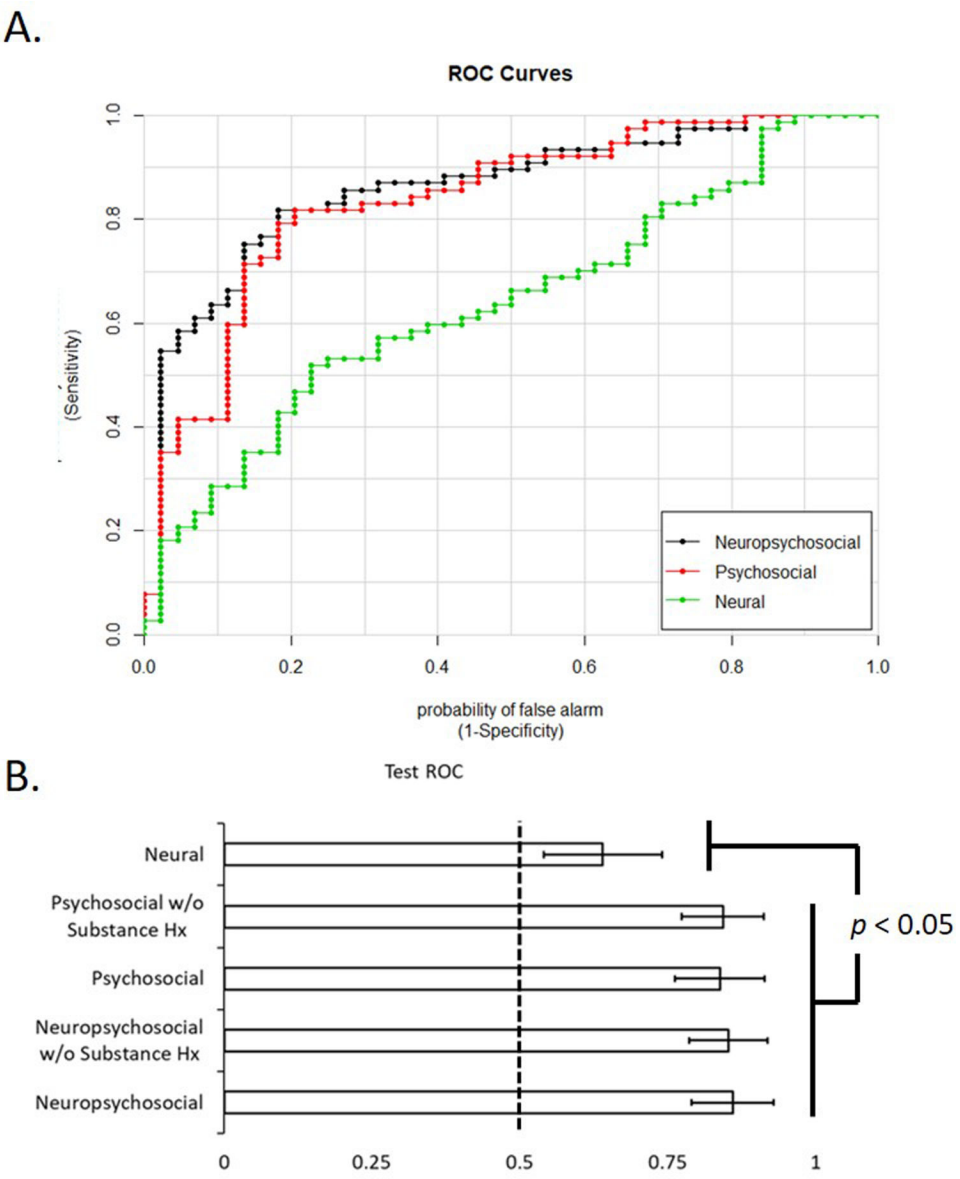

Figure 3. Performance of Neuropsychosocial, Psychosocial, and Neural Models.

Panel A depicts the receiver operating characteristic plot for the ensemble models generated from each data source when applied to the test sample. Panel B depicts the area under the curve for the plots and the error bars represent the 95% confidence interval. Panel B also includes the psychosocial datasets that excluded substance use history to determine how much influence it had on model performance. The dashed line represents chance performance, so if the error bars do not contain the dashed line, they can be considered to perform better than chance with a probability of $p < 0.05$.

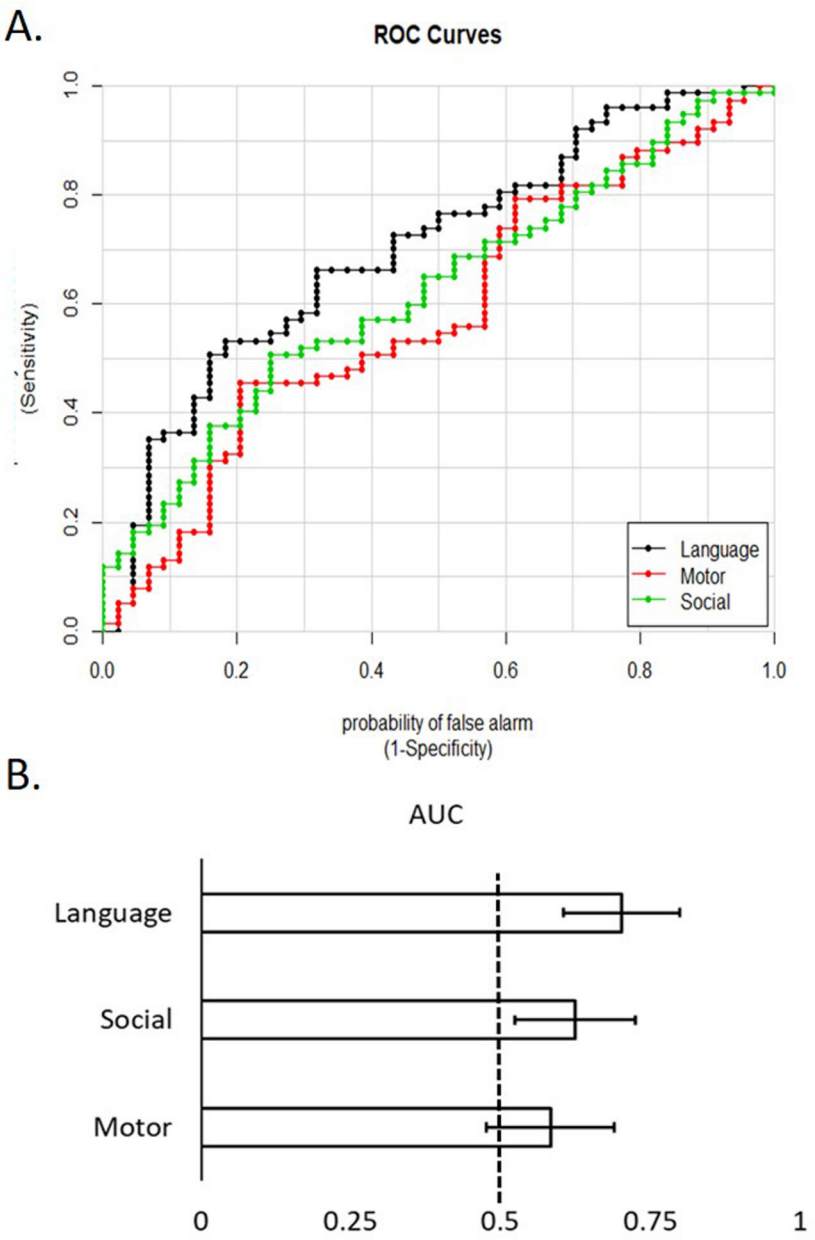

Figure 4. Performance of Language, Motor, and Social fMRI models.

Panel A depicts the receiver operating characteristic plot for the ensemble models generated from each data source when applied to the test sample. Panel B depicts the area under the curve for the plots and the error bars represent the 95% confidence interval. The dashed line represents chance performance, so if the error bars do not contain the dashed line, they can be considered to perform better than chance with a probability of $p < 0.05$.

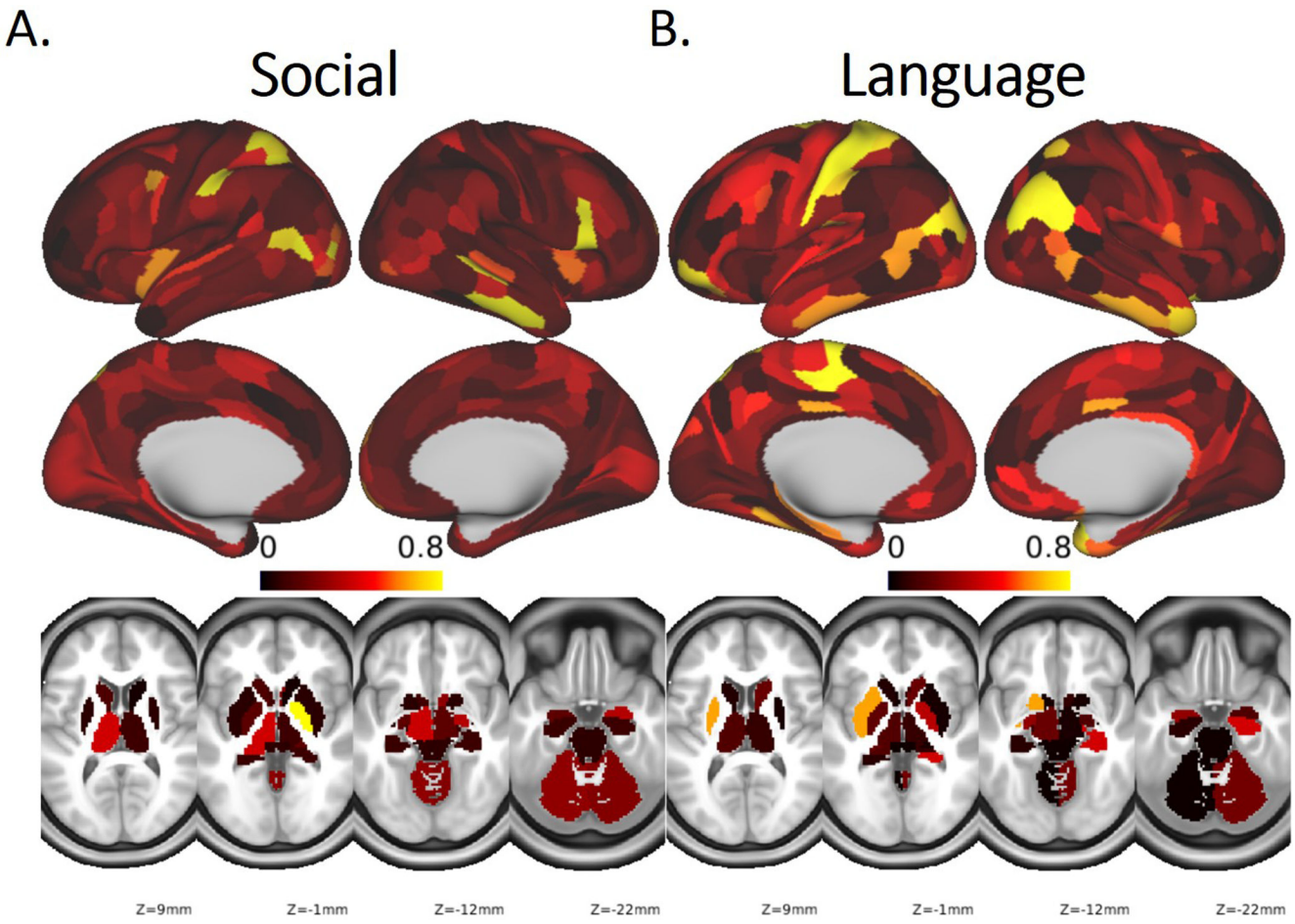

Figure 5. Neural Contributions to Classification.

Variable importance scores were generated for the models of the social and language fMRI tasks. Each region received a score. Higher scores indicate greater importance in the model's classification scheme. Brighter colors depict greater variable importance, with yellow being the most important. Panel A depicts scores from the social task and Panel B depicts scores from the language task. Approximately 10 regions per task contributed to classification, suggesting that neural models of risky drinking may need to incorporate complex models to characterize substance use problems accurately.

Table 1.

Sample characteristics

	Binge N=177	**Non-Binge N=309**	
	N, %	**N, %**	X^2
Male	128, 72.3	101, 32.7	70.9 **
Monozygotic Twin	39, 22.0	96, 31.1	4.6 *
Dizygotic Twin	44, 24.9	71, 23.0	0.2
White	147, 83.1	214, 69.3	11.2 **
Black	15, 8.5	61, 19.7	10.8 **
Hispanic	13, 7.3	27, 8.7	0.3
Alcohol Dependence	30, 16.9	0, 0	---
Alcohol Abuse	62, 35.0	0, 0	---
	(mean ± SD)	**(mean ± SD)**	***t*-statistic**
Age (years)	27.9 ± 3.4	29.4 ± 3.9	4.27 **
Body Mass Index	26.8 ± 4.1	26.1 ± 5.4	1.49
Income (per year)	$47k ± $20k	$51k ± $20k	1.01
Education (Years)	14.6 ± 1.9	15.0 ± 1.8	2.31 *
Drinks over last 7 days	13.0 ± 11.1	1.5 ± 2.7	11.50 **
Tobacco uses over last 7 days	18.3 ± 35.8	5.0 ± 18.9	13.35 **

* denotes $p < 0.05$

** denotes $p < 0.001$

Table 2.

Comparison of Model Performance.

Model performance was assessed primarily with the area under the receiver operating characteristic curve, but brier score, accuracy, and sensitivity are also reported. Significance was defined as when the lower bound of the 95% confidence interval was above 0.50. For the fMRI tasks, we only examined performance in the tasks that performed better than chance in the training sample.

Model	Area under ROC	95%CI Lower	95%CI upper	Brier Score	Accuracy	Sensitivity	Specificity
Neuropsychosocial	0.86	0.79	0.93	0.15	0.80	0.70	0.86
Neuropsychosocial w/o Substance Hx	0.85	0.79	0.92	0.15	0.77	0.74	0.78
Psychosocial	0.84	0.76	0.91	0.16	0.79	0.80	0.79
Psychosocial w/o Substance Hx	0.84	0.77	0.91	0.16	0.78	0.68	0.83
Neural	0.64	0.54	0.74	0.22	0.64	0.30	0.83
Neural domain comparisons							
Language	0.70	0.61	0.80	0.20	0.66	0.32	0.86
Motor	0.58	0.48	0.69	0.23	0.63	0.09	0.94
Social	0.62	0.52	0.73	0.22	0.63	0.16	0.90