Predicting individual risk tolerance from brain anatomy with interpretable machine learning models

Francis Fan*, Peng Quan#, Tianxin Mao#, Hengyi Rao#

*Harriton High School, 600 North Ithan Avenue, Rosemont, PA 19010, USA

#Center for Functional Neuroimaging, Department of Neurology, University of Pennsylvania, Room D502, Richards Medical Research Building, 3700 Hamilton Walk, Philadelphia, PA 19104, USA

Abstract

We used interpretable machine learning models and investigated the predictive power of brain anatomy for individual risk tolerance in 118 healthy participants with structural MRI scans. Individual risk-taking behavior was measured by the Balloon Analogue Risk Task (BART) and risk tolerance was characterized by the extended Expectancy Valence model. Multiple machine learning models with cross-validation were adopted based on regional gray matter volumes and Shapley values were used to identify the most informative brain regions. The findings demonstrated that the left cerebellum and right posterior parietal cortex were the most important brain regions, solidifying findings of prior research.

1. Introduction

Understanding which regions of the brain play the largest roles in risk-taking tendencies can have far-reaching consequences, from advancing addiction research to increasing understandings in human behavior (1-4). Prior work has shown that there is substantial association between individual human decision-making and neuroanatomy, particularly gray matter volume (GMV) of brain regions such as the amygdala, posterior parietal cortex (PPC), and a left cerebellar region (2-4). However, most studies used univariate statistical methods to analyze these separate regions in a group level. Here, we used interpretable machine learning methods to build and explain predictive models for predicting the risk-taking behavior measures, specifically risk tolerance, based on gray matter volumes of brain regions associated with risk-taking behavior.

2. Methods

We adopted interpretable machine learning methods to build predictive models for predicting the risk-taking behavior measures based on gray matter volumes of brain regions associated with risk-taking behavior in a sample of 118 healthy participants with structural MRI scans and risk-taking behavior/risk tolerance measures during the Balloon Analogue Risk Task (BART), a well-established sequential risky decision paradigm. We explained the machine learning models with Shapley values (5), a game theory based method (6), to identify the most informative brain regions predictive of risk-taking behavior.

2.1. Participants

A cohort of 118 heathy adults were recruited and all the participants were right-handed, had a normal or corrected-to-normal vision, and free of neurological or psychiatric issues. The study procedure was approved by the University of Pennsylvania institutional review board. Written informed consent was obtained from all participants.

2.2. BART Procedure and Image acquisition

A MRI-compatible BART paradigm was adopted to measure individual risk-taking behavior within the MRI scan chamber (7). Specifically, the BART requires participants to inflate a virtual balloon that could either grow larger or explode. For each balloon, the participant may continue or discontinue inflating the balloon by pressing two buttons. Larger balloons are always associated with greater risk of explosion and increased monetary rewards. Participants may make multiple inflation attempts to try and maximize their monetary rewards. If the balloon explodes upon inflation, they lose the amount waged on the current balloon from their total winnings. If participants decide to stop inflating the balloon, they may collect the wager for the current balloon. The timing of inflation is controlled by a cue and participants press a button to continue or discontinue inflation only when the color of the cue is green.

In the BART, the outcome for each trial was immediately provided to participants once they collected the wager or the balloon exploded. Each participant's BART performance was quantified by a risk tolerance measure, γ. A detailed description of the BART and the risk tolerance measure was presented in our previous study (2).

MRI scans were performed using a 3T Siemens Trio scanner (Siemens Medical Systems, Erlangen, Germany). High-resolution anatomic images were obtained using a T1-weighted 3D Magnetization Prepared Rapid Acquisition Gradient Echo (MPRAGE) sequence with repetition time (TR) = 1620 ms, echo time (TE) = 3.09 ms, flip angle (FA) = 15°, 176 contiguous slices, matrix size = 192 × 256, voxel size = 0.98 × 0.98 × 1.0 mm^3. During the scan, the participants were instructed to keep their eyes open and fixate on a cross mark in the center of the screen. All the MRI scans were processed using sing CAT12 (www.neuro.uni-jena.de) and SPM12 toolbox (www.fil.ion.ucl.ac.uk), implemented in MATLAB 2016, as detained in our previous study (2). Gray matter volume (GMV) measures were computed for all regions of interest, including bilateral amygdala, PPC, and a left cerebellar region (2-4).

2.3. Interpretable machine learning

We used multiple machine learning methods, including random forests (8), support vector machines (9), multivariable linear regression, and a generalized additive model (10), to build predictive models for predicting individual risk tolerance measures based on GMV measures of brain regions associated with risk-taking behavior. The brain regions of interest included bilateral amygdala, posterior parietal cortex (PPC), and a left cerebellar region, all identified in prior studies (2-4). Their GMV measures were used as features to predict individual risk tolerance measures, with age, sex, and total GMV measure as covariates. All the machine learning methods were implemented using Scikit-learn (11).

All the machine learning models were evaluated using the same 10-fold cross-validation by randomly splitting the whole cohort into 10 non-overlapping folds and using each of them as a testing subset and the remaining as a training subset. This procedure was repeated 1000 times in order to obtain a robust estimation of the prediction performance. We adopted Pearson correlation between the predicted and measured risk tolerance measures to evaluate the prediction performance.

We explained the machine learning models with Shapley values, a game theory based method, to identify the most informative brain regions predictive of the risk-taking behavior. We adopted SHAP (SHapley Additive exPlanations), a python package (12), to explain the machine learning models and quantify the importance of each of the brain regions of interest for the prediction. As a game theoretic approach, it treats each feature as a player and quantifies each player's importance and contribution to the overall prediction using the classic Shapley values (5, 6).

3. Results

Characteristics of the participants are summarized in Table 1, including sex, age, total gray matter volume, and risk tolerance score.

Table 1. Participant characteristics

Total number of participants: n=118	
Sex	Male: 62, female: 56
Age (years)	Mean: 29.7, Median: 27.0, Range: 21-50
Total gray matter volume (mm^3)	Mean: 644.5, Median: 643.8, Range: 474.0-822.2
Risk tolerance measure (γ)	Mean: 0.46, Median: 0.44, Range: 0.27-0.71,

Table 2 summarizes performance of prediction models built using different machine learning methods, including random forests (RF) (8), support vector regression (SVR) (9), multivariable linear regression (MLR), and a generalized additive model (EBM) (10). Each method's performance was estimated on the same 1000 runs of the 10-fold cross-validation. These results indicated that all the machine learning models successfully predicted risk tolerance measures that were significantly correlated with the measured values. Among the machine learning methods under investigation, random forests achieved the best performance.

Table 2. Prediction performance of different machine learning methods under comparison, including random forests (RF), multivariable linear regression (MLR), support vector regression (SVR), and a generalized additive model (EBM). The performance was measured by mean of Pearson correlation coefficients of 1000 runs of the 10-fold cross-validation, and their p-values.

Methods	**RF**	**MLR**	**SVR**	**EBM**
Mean of correlation coefficients	0.3924	0.3586	0.3261	0.3043
Mean of p-values	<0.0001	0.0001	0.0004	0.0013
Maximum of p-values	0.0005	0.0016	0.0021	0.0147

In order to explain the machine learning models and identify the most important brain regions for the prediction, we computed Shapley values of the RF models for each of the individual brain regions of interest as well as age, sex, total gray matter volume. As shown in Figure 1, the left cerebellar region and the right PPC were identified as the most important brain regions, solidifying findings of prior research.

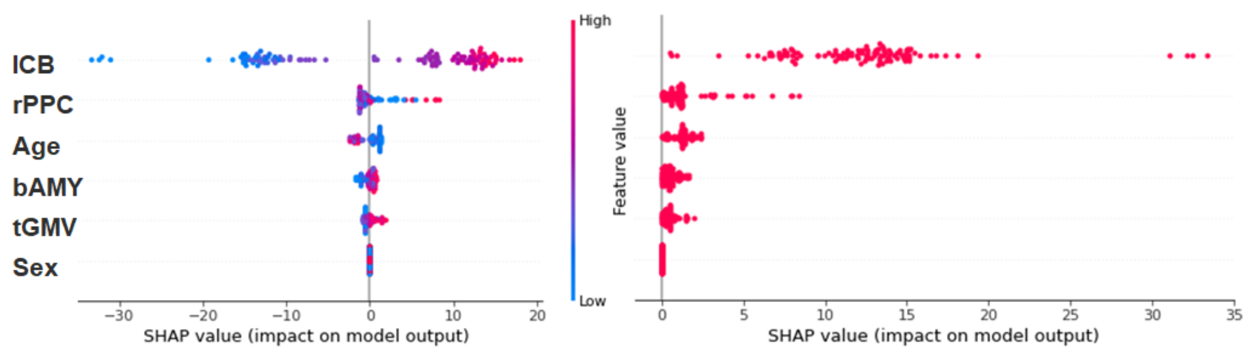

Figure 1. Shapley values (Left: original value; Right: absolution value) of GMVs of brain regions of interest, age, gender, and total GMV. GMV = grey matter volume; lCB = left cerebellum; rPPC = right posterior parietal cortex; bAMY = bilateral amygdala; tGMV = total GMV of the brain.

4. Conclusions

Our study has demonstrated that machine learning models can successfully predict individual risk tolerance based on brain anatomy measures. Specifically, the left cerebellum and the right PPC were identified as the most indicative brain regions for the prediction, solidifying findings of prior research. These findings demonstrate the utility of interpretable machine learning models as a quantitative tool to estimate individual behavior based on neuroanatomy measures.

The work is not being, or has been, submitted for publication or presentation anywhere.

We are currently working on building machine learning models on the whole brain GMV measures. We will report all the results in the final manuscript. To facilitate reproducible research, we will make our code publicly available on GitHub once the paper is accepted.

References

1. E. U. Weber, A.-R. Blais, and N. E. Betz, "A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors," *Journal of Behavioral Decision Making* **15**(4), 263-290 (2002).
2. P. Quan et al., "Cerebellum anatomy predicts individual risk-taking behavior and risk tolerance," *NeuroImage* **254**(119148 (2022).
3. S. Gilaie-Dotan et al., "Neuroanatomy Predicts Individual Risk Attitudes," *The Journal of Neuroscience* **34**(37), 12394-12401 (2014).
4. W. H. Jung et al., "Amygdala Functional and Structural Connectivity Predicts Individual Risk Tolerance," *Neuron* **98**(2), 394-404.e394 (2018).
5. S. M. Lundberg, and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 4768–4777, Curran Associates Inc., Long Beach, California, USA (2017).
6. A. E. Roth, *The Shapley Value: Essays in Honor of Lloyd S. Shapley*, Cambridge University Press, Cambridge (1988).
7. H. Rao et al., "Neural correlates of voluntary and involuntary risk taking in the human brain: An fMRI Study of the Balloon Analog Risk Task (BART)," *NeuroImage* **42**(2), 902-910 (2008).
8. L. Breiman, "Random Forests," *Machine Learning* **45**(1), 5-32 (2001).
9. C.-C. Chang, and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.* **2**(3), Article 27 (2011).
10. Y. Lou et al., "Accurate intelligible models with pairwise interactions," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 623–631, Association for Computing Machinery, Chicago, Illinois, USA (2013).
11. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research* **12**(2825--2830 (2011).
12. SHAP Team, "SHAP: https://github.com/slundberg/shap," (2022).