GROWUPGROWOLD
HUMAN DEVELOPMENT, BIRTH TO DEATH

Moffitt & Caspi
Genes Environment Health Behavior

Concept Paper Form

Provisional Paper Title: Multimodal brain-based predictive model for cognitive changes in middle-aged adults
Proposing Author: Narun Pat
Author's Email: narun.pat@otago.ac.nz
P.I. Sponsor: Ahmad Hariri (if the proposing author is a student or colleague of an original PI)
Today's Date: 11/11/2021

Please describe your proposal in 2-3 pages with sufficient detail for helpful review.

Objective of the study:

Research has long recognized the changes in cognitive abilities as a normal aging process, even with the absence of disease[1–3]. On the one hand, crystalized abilities that involve past learning experiences (such as, verbal comprehension) is usually intact, if not improved overtime. On the other hand, fluid cognitive abilities that involve new learning and information processing (such as, processing speed, working memory and perceptual reasoning) is often worsen as people age. These trajectories can be seen in the young adult onward[2]. More importantly, individuals vary considerably in these trajectories[4]. Understanding individual differences in the aged-related changes in cognitive abilities is important as (1) these changes are often associated with changes in day-to-day functioning and (2) these changes can provide a basis of comparison between normal aging and disease states[2,3]. Here, capitalizing on high-quality longitudinal data from the DMHDS, we will use neuroimaging and predictive modeling approaches to elucidate the extent to which the changes in cognitive abilities from the young adult to the middle age are reflected in the brain functions and structure.

Substantial efforts in neuroimaging have been paid to study aged-related changes in cognitive abilities. However, past studies often suffer from many methodological issues[3]. Fortunately, many of these issues can be addressed by the DMHDS. First, studies usually have trouble with selection bias. They often have difficulty recruiting participants who are either very healthy (who are usually too busy) or very unhealthy as well as those who have limited financial and social support[3,5,6]. The DMHDS, as a population-based study, does not have this issue, given that the DMHDS includes 91% of eligible births who were born in the same year in Dunedin, NZ[7]. Second, many studies are cross-sectional[8]. Examining age-related changes in cross-sectional studies can be confounded by the cohort effects (i.e.,

differences in life experiences due to the year of birth). Moreover, a cross-sectional design makes it difficult to estimate within-person changes in cognitive abilities between time points since there is only one time point. The DMHDS is a longitudinal study, and thus does not suffer from these drawbacks. More specifically, the current DMHDS will allow us to examine the longitudinal changes in cognitive abilities between age 38 and 45. Third, many longitudinal studies suffer from attrition over time, leaving them with those healthiest, wealthiest and most cognitive capable in the cohort[9]. The DMHDS is an exception where 94% of the alive participants still took part in the study after a number of assessments from prenatal to 45 years old[7]. Accordingly, the DMHDS provides a unique opportunity to address these problematic issues that have prevented the understanding of age-related changes in cognitive abilities at the neural level.

Nonetheless, establishing a cognition-brain relationship from neuroimaging data has proven to be challenging[21] in terms of predictability, reliability, and interpretability. We, however, believe that these challenges can be addressed by modern predictive modeling approaches to neuroimaging[10]. First, prediction is an ability to estimate changes in cognitive abilities of out-of-sample individuals (i.e., not part of the model-building process) based on their brain data[11]. Although it is still less common for studies to predict the *changes* in cognitive abilities overtime, many studies have used predictive modelling (also known as multivariate/machine learning methods) to draw information across brain indices in order to cross-sectionally predict individual differences in cognitive abilities from Magnetic Resonance Imaging (MRI)[10]. These studies often implement MRI data of a single modality, such as structural MRI (sMRI), resting-state functional connectivity (rs-FC) and task-based functional MRI (tfMRI). sMRI reflects brain volume and morphology[26,27]. A recent large-scale predictive competition[12] shows a weak relationship between sMRI and cognitive abilities at out-of-sample *r* around .03-.15. rs-FC reflects intrinsic functional connectivity between different brain areas during rest. Studies generally find better prediction of cognitive abilities from rs-FC, compared to sMRI, at *r* around .2 to .4[13–15]. tfMRI reflects the changes in BOLD signal induced by events embedded in tasks. Recent evidence seems to suggest that tfMRI from certain tasks provide superior prediction for cognitive abilities at *r* above .4[16]. More recently, as opposed to relying on each single modality, research has started to combine data across MRI modalities via a machine-learning technique called stacking. Stacked models, for instance, enhance prediction of participants' age[17] and cognitive abilities[14,18] over and above predictive models based on single modalities. Accordingly, combining multiple modalities via stacking should also enhance prediction for the changes in cognitive ability in middle-aged adults, especially if each MRI modality provides a separatable, but overlapping, contribution.

Second, reliability, or more specifically test-retest reliability, implies that a brain-based predictive model should provide a similar predicted value of the changes in cognitive ability across measurement times[19]. Researchers often quantify reliability using intraclass correlation (ICC), whereas ICC < .40 reflects poor reliability, and ICC > .75 reflects excellent reliability[20]. Similar to prediction, different modalities demonstrate different levels of reliability. tfMRI in particular has recently come under intense scrutiny for its low reliability[21]. Elliot and colleagues[21] examined ICC of tfMRI at different regions using the tasks from the DMHDS and showed poor ICC (<.4) across the regions and tasks. This is sharply contradicting the ICC of sMRI, which is at the excellent range (>.75). Fortunately, recent studies have started to show that drawing information across brain regions (as

opposed to relying on one region at a time) in predictive modelling markedly boosts reliability of tfMRI[22]. More recently, we demonstrated that combining information across tasks via stacked modelling can further enhance reliability of tfMRI, resulting in the excellent range of ICC[18]. Accordingly, stacked models should improve not only prediction, but also reliability, of the predictive models for the changes in cognitive abilities. Fortunately, the DMHDS provides data from 20 participants who were scanned twice at the age of 45. This will allow us to examine the reliability for our predictive models via ICC.

Third, interpretability is the extent to which we are able to understand contribution from each brain feature in the predictive models[23]. Ability to interpret models is the key to gain neurobiological insights into the changes in cognitive abilities in the middle age. Nonetheless, not all predictive modelling algorithms are easy to interpret. Some algorithms, especially those that assume linearity and additivity, such as Elastic Net[24] and support vector machine (SVM) with a linear kernel[25], are easy to interpret—researchers can simply examine the algorithm's coefficient (i.e., weight) of each brain feature. It is, however, harder to interpret contribution from each brain feature with other algorithms that allow for nonlinearity and interaction among features. Some examples of these algorithms are SVM with non-linear kernels (e.g., the radial basis function and polynomial kernel[26]) and ensemble learning (e.g., random forest[27] and XGBoost[28]). This problem is known as the accuracy-interpretability trade off in machine-learning[29]. Fortunately, modern machine-learning has provided model-agnostic techniques to assist with interpreting complex models[23], such as SHapley Additive exPlanations (SHAP)[30], Accumulated Local Effects (ALE)[31] and feature interaction[32]. These techniques allow us to interpret and visualise feature importance, (non)-linearity and interactivity, respectively. Accordingly, we will apply different algorithms (including Elastic Net, SVM and ensemble learning) and select the one with the highest predictability and reliability. We will then use appropriate techniques to interpret the final models. In keeping with the parieto-frontal integration theory of intelligence[33] and a recent meta-analysis[34], we expect that regions in the frontoparietal as well as dorsal attention and default-model networks to contribute highly in predicting the age-related changes in cognitive abilities.

Our overarching goal is to enhance our neurobiological understanding of individual differences in the age-related changes in cognitive abilities. We will achieve this by developing a predictable, reliable, and interpretable model of the individual differences from multimodal MRI (sMRI, rs-FC and tfMRI from different tasks). More specifically, our predictive models will be tuned to predict changes in cognitive abilities of both crystalized and fluid abilities. Similar to our recent predictive modelling work[18], we will draw information across the whole brain from each MRI modality via different algorithms (Elastic Net[24], SVM[25] and ensemble learning[27]), and then combine data across modalities via stacking. Once we find a predictive and reliable model, we will interpret them using appropriate techniques (e.g., relying on coefficients, SHAP[30], ALE[31] and/or feature interaction[32]).

Data analysis methods:

Target Variables:

Cognitive abilities will be based on (1) cognitive assessment scores during the time of MRI scanning (i.e., age 45) and (2) the differences in cognitive assessment scores between age 38 and 45. The first is for prediction for the current cognitive abilities, while the latter is for prediction for the changes in cognitive abilities as a function of age.

In order to capture both crystalized (verbal comprehension) and fluid (processing speed, working memory and perceptual reasoning) cognitive abilities, we will use the WAIS-IV[35]. Specifically, we will used (1) information, similarities, and vocabulary scores for verbal comprehension, (2) symbol search for processing speed, (3) arithmetic and digit span for working memory and (4) block design, matrix reasoning and picture completion for perceptual reasoning.

Feature Variables:

Multimodal Neuroimaging: We will use neuroimaging data from task-based functional MRI (tfMRI), resting-state fMRI connectivity (rs-FC) and structural MRI (sMRI) as the input features for our model:

1) For task-based functional MRI (tfMRI), we will focus on whole-brain parcel-wise contrast values. We will parcellate tfMRI into 379 regions using Glasser's cortical atlas[36] and Freesurfer's subcortical segmentation[37]. We will then compute averaged general-linear model contrasts between experimental vs. control conditions for each tfMRI task as different modalities. For the emotion task, we will focus on the "Faces > Shapes" contrast[21]. For the reward task, we will focus on the "Gain Anticipation > Neutral Anticipation" contrast[21]. For the executive function Stroop task, we will focus on the "Incongruent > Congruent" contrast[21]. For the episodic memory task, we will focus on the "Encoding > Distractor" contrast[21].

2) For resting-state fMRI connectivity (rs-FC), we will quantify connectivity strength during rs-FC using a seed-based, correlational approach on parcellated regions. Similar to tfMRI, we will parcellate tfMRI into 379 regions using Glasser's cortical atlas[36] and Freesurfer's subcortical segmentation[37]. We will compute rs-FC connectivity-strength indices using Fisher r to z transformation.

3) For structural MRI (sMRI), we will separate sMRI data into four different modalities: cortical thickness, cortical surface area, subcortical volume and total brain volume. For cortical thickness and cortical surface area, we used Destrieux parcellation (148 ROIs)[38,39]. As for subcortical volume, we used FreeSurfer's subcortical segmentation (19 gray matter ROIs)[37]. As for total brain volume, we included five features calculated by FreeSurfer: estimated intra-cranial volume, total cortical gray matter volume, total cortical white matter volume, total subcortical gray matter volume and ratio of brain segmentation volume to estimated total intracranial volume.

Model Building:

We will use a nested, stacking cross-validation (see Figure 1). In each of the cross-validation (CV) “outer” loops, one of the eight folds will be held-out. The rest will be further split into 60% and 40% for the first- and second-layer training layers, respectively. Within the CV “inner” loops, we will separately fit the first-layer data from each modality to predict a target variable. Here we will apply a five-fold CV to tune hyperparameters of the models. This stage will allow us to create modality-specific models (sMRI, rs-FC and tfMRI from each of the three tasks).

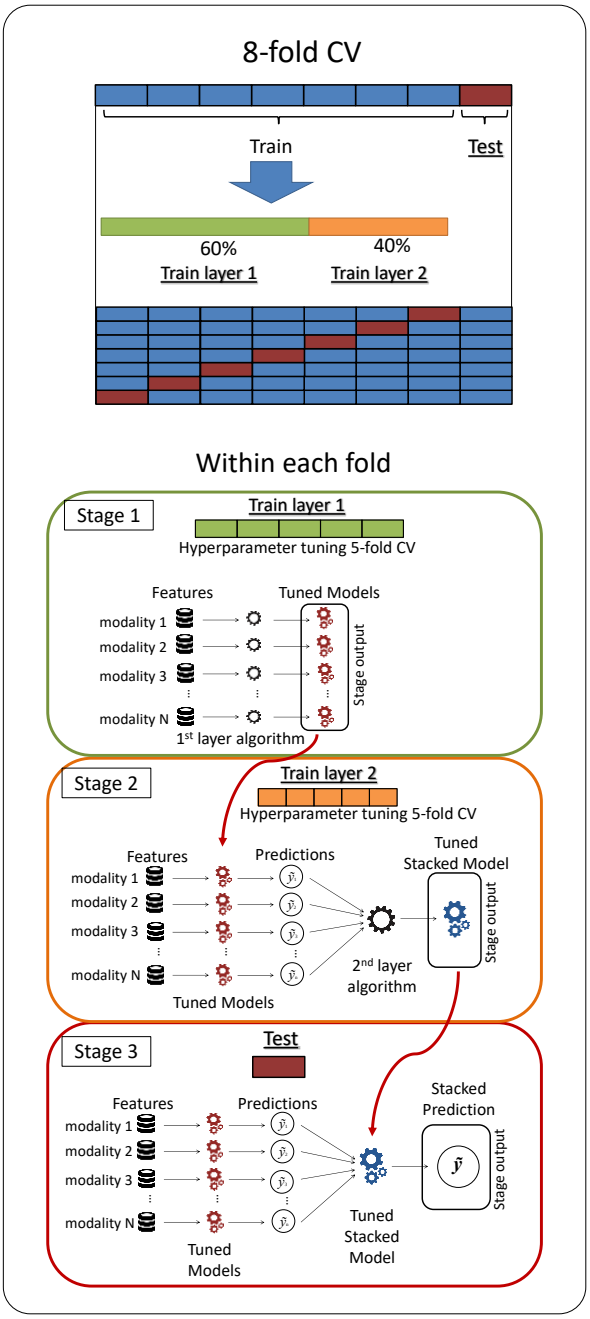

Figure 1. Model building pipeline

Model Algorithms:

For both first-layer and second-layer algorithms, we will compare the prediction and reliability among the following six algorithms from three main supervised-learning families:

1. Linear regression:
    a. Elastic Net[24]
2. Support vector machine (SVM):
    a. SVM with a linear kernel[25]
    b. SVM with a radial basis function kernel[26]
    c. SVM with a polynomial kernel[26]
3. Ensembles:
    a. random forest[27]
    b. XGBoost[28]

Different algorithms are sensitive to different pattern of the data (e.g., whether the brain features have a linear or non-linear relationship with cognitive abilities and whether there are interactions among the brain features)[29]. Thus, evaluating different algorithms will ensure that we are able to have the best performing algorithms.

Note some algorithms (e.g., random forest) have an addition advantage (apart from prediction and reliability) in dealing with missing values. A recent study in aging population[17], for instance, used random forest as the second-layer algorithm along with imputation, allowing them to keep missing values due to MRI artifact in certain MRI modalities in the final model. The ability to keep more data in the model using this so-called opportunistic stacking may outweigh potential poorer prediction and/or reliability.

Model Prediction:

We will evaluate models' prediction using the eight held-out folds across the outer CV loops via four measures[40]:

1. Pearson's *r* is defined as $\frac{cov(y,\hat{y})}{\sigma_y \sigma_{\hat{y}}}$

2. coefficient of determination (R^2) is defined as $1 - \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2}$
3. mean squared error (MSE) is defined as $\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
4. mean absolute error (MAE) is defined as $\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$

where cov is the covariance, σ is the standard deviation, y is the observed value and ŷ is the predicted value, where $\overline{y}$ is the mean of the observed value.

Model Reliability:

We will use two types of intraclass correlation (ICC) [41] that are commonly used in MRI studies[19]:

1. ICC(2,1) is defined as $\frac{MS_p - MS_e}{MS_p + (k-1)MS_e + \frac{k}{n}(MS_t - MS_e)}$
2. ICC (3,1) is defined as $\frac{MS_p - MS_e}{MS_p + (k-1)MS_e}$

where MS_p is mean square for participants, MS_e is mean square for error, MS_t is mean square for time points (i.e., measurements), n is the number of participants, k is the number of time points.

Model Interpretability (i.e., feature importance):

Examining feature importance of the layer-one model will allow us to demonstrate which of the brain indices contribute highly to the prediction of the modality-specific models. Similarly, examining feature importance of the layer-two model will allow us to demonstrate which of the modalities contribute highly to the prediction of the stacked models. For linear algorithms (Elastic Net[24] and SVM with a linear kernel[25]), we will examine their coefficient weight for each brain index. For other algorithms, we will use SHAP[30], ALE[31] and feature interaction[32].

More specifically, once we identify which of the modalities contribute highly to the prediction of the stacked model (based on the feature importance indices, such as coefficient weight and/or SHAP), we will then examine the brain features of the top-performing modality-specific models that contributed highly to the prediction (again, based on the feature importance indices). To support the parieto-frontal integration theory of intelligence[33], we expect to see the areas in the frontoparietal network as top-performing brain features across top-performing MRI modalities in predicting the changes in cognitive abilities.

Construct validity:

To ensure the construct validity of our predictive model for the changes in cognitive abilities, we will also test the relationship between quality of sleep and predicted value of the final model. Studies show the relationship between poor cognitive abilities and sleep in aging individuals[44], and thus our brain-based model for the changes in cognitive abilities should also be varied as a function of the quality of sleep.

Variables needed at which ages:

DMHDS Age 38 variables:

1. Cognitive assessment:
    - All WAIS-IV scores, including information, similarities, vocabulary, symbol search, arithmetic, digit span, block design, matrix reasoning and picture completion
2. Sleep quality

DMHDS Age 45 variables:

1) Cognitive assessment
    - All WAIS-IV scores, including information, similarities, vocabulary, symbol search, arithmetic, digit span, block design, matrix reasoning and picture completion
2) Sleep quality
3) Neuroimaging:
    - Structural MRI
    - fMRI time course for resting state
    - fMRI time course for the Emotion task
    - fMRI time course for the Reward task
    - fMRI time course for the Stroop task
    - fMRI time course for the Memory task

Significance of the Study (for theory, research methods or clinical practice):

We aim to better understand the neural basis of aged-related changes in cognitive abilities from young adult to middle age. Our brain-based predictive models built from the population-based DMHDS will provide a baseline for normal aging against which abnormal aging can be compared. Given the use of high-quality longitudinal data of DMHDS, our predictive model will not suffer from issues commonly found in other aging studies, such as the selection bias, cohort effect and attrition. Moreover, our predictive modelling approach will maximize predictability of the models, while still allowing for neurobiological insights through their interpretability. Thus, according to the parieto-frontal integration theory of intelligence[33] and a recent meta-analysis[34], we will examine the role of areas in the frontoparietal network across different MRI modalities in predicting the changes in cognitive abilities. Next, we will also demonstrate which of the modalities capture cognitive ability in middle-age population, and which do not. This information will move the field forward by allowing researchers to focus on informative modalities. Accordingly, using the rich MRI dataset from the DMHDS will provide a unique opportunity for us to enhance the neurobiological understanding aged-related changes in cognitive abilities.

References cited:

1. Horn JL, Cattell RB. Age differences in fluid and crystallized intelligence. *Acta Psychologica*. 1967;26:107-129. doi:10.1016/0001-6918(67)90011-X

2. Salthouse TA. Trajectories of normal cognitive aging. *Psychology and Aging*. 2019;34(1):17-24. doi:10.1037/pag0000288

3. Harada CN, Natelson Love MC, Triebel KL. Normal Cognitive Aging. *Clinics in Geriatric Medicine*. 2013;29(4):737-752. doi:10.1016/j.cger.2013.07.002

4. Wisdom NM, Mignogna J, Collins RL. Variability in Wechsler Adult Intelligence Scale-IV Subtest Performance Across Age. *Archives of Clinical Neuropsychology*. 2012;27(4):389-397. doi:10.1093/arclin/acs041

5. Ford JG, Howerton MW, Lai GY, et al. Barriers to recruiting underrepresented populations to cancer clinical trials: A systematic review. *Cancer*. 2008;112(2):228-242. doi:10.1002/cncr.23157

6. Minder CE, Müller T, Gillmann G, Beck JC, Stuck AE. Subgroups of Refusers in a Disability Prevention Trial in Older Adults: Baseline and Follow-Up Analysis. *Am J Public Health*. 2002;92(3):445-450. doi:10.2105/AJPH.92.3.445

7. Poulton R, Moffitt TE, Silva PA. The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(5):679-693. doi:10.1007/s00127-015-1048-8

8. Taylor JR, Williams N, Cusack R, et al. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *NeuroImage*. 2017;144:262-269. doi:10.1016/j.neuroimage.2015.09.018

9. Van Beijsterveldt CEM, van Boxtel MPJ, Bosma H, Houx PJ, Buntinx F, Jolles J. Predictors of attrition in a longitudinal cognitive aging study:: The Maastricht Aging Study (MAAS). *Journal of Clinical Epidemiology*. 2002;55(3):216-223. doi:10.1016/S0895-4356(01)00473-5

10. Sui J, Jiang R, Bustillo J, Calhoun V. Neuroimaging-based Individualized Prediction of Cognition and Behavior for Mental Disorders and Health: Methods and Promises. *Biological Psychiatry*. 2020;88(11):818-828. doi:10.1016/j.biopsych.2020.02.016

11. Yarkoni T, Westfall J. Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspect Psychol Sci*. 2017;12(6):1100-1122. doi:10.1177/1745691617693393

12. Mihalik A, Brudfors M, Robu M, et al. ABCD Neurocognitive Prediction Challenge 2019: Predicting Individual Fluid Intelligence Scores from Structural MRI Using Probabilistic Segmentation and Kernel Ridge Regression. In: Pohl KM, Thompson WK, Adeli E, Linguraru MG, eds. *Adolescent Brain Cognitive Development Neurocognitive Prediction*. Lecture Notes in

Computer Science. Springer International Publishing; 2019:133-142. doi:10.1007/978-3-030-31901-4_16

13. Sripada C, Angstadt M, Rutherford S, Taxali A, Shedden K. Toward a “treadmill test” for cognition: Improved prediction of general cognitive ability from the task activated brain. *Human Brain Mapping*. 2020;41(12):3186-3197. doi:10.1002/hbm.25007

14. Rasero J, Sentis AI, Yeh FC, Verstynen T. Integrating across neuroimaging modalities boosts prediction accuracy of cognitive ability. *PLOS Computational Biology*. 2021;17(3):e1008347. doi:10.1371/journal.pcbi.1008347

15. Dubois J, Galdi P, Paul LK, Adolphs R. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2018;373(1756):20170284. doi:10.1098/rstb.2017.0284

16. Barch DM, Burgess GC, Harms MP, et al. Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage*. 2013;80:169-189. doi:10.1016/j.neuroimage.2013.05.033

17. Engemann DA, Kozynets O, Sabbagh D, et al. Combining magnetoencephalography with magnetic resonance imaging enhances learning of surrogate-biomarkers. Shackman A, de Lange FP, Tsvetanov K, Trujillo-Barreto N, eds. *eLife*. 2020;9:e54055. doi:10.7554/eLife.54055

18. Tetereva A, Li J, Deng J, Stringaris A, Pat N. *Integrating Task-Based Functional MRI Across Tasks Markedly Boosts Prediction and Reliability of Brain-Cognition Relationship*. bioRxiv; 2021. doi:10.1101/2021.10.31.466638

19. Noble S, Scheinost D, Constable RT. A guide to the measurement and interpretation of fMRI test-retest reliability. *Current Opinion in Behavioral Sciences*. 2021;40:27-32. doi:10.1016/j.cobeha.2020.12.012

20. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic*. 1981;86(2):127-137.

21. Elliott ML, Knodt AR, Ireland D, et al. What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis. *Psychol Sci*. 2020;31(7):792-806. doi:10.1177/0956797620916786

22. Kragel PA, Han X, Kraynak TE, Gianaros PJ, Wager TD. Functional MRI Can Be Highly Reliable, but It Depends on What You Measure: A Commentary on Elliott et al. (2020). *Psychol Sci*. 2021;32(4):622-626. doi:10.1177/0956797621989730

23. Molnar C. *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*.; 2019. https://christophm.github.io/interpretable-ml-book/

24. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2005;67(2):301-320. doi:10.1111/j.1467-9868.2005.00503.x

25. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273-297. doi:10.1007/BF00994018

26. Chang YW, Hsieh CJ, Chang KW, Ringgaard M, Lin CJ. Training and Testing Low-degree Polynomial Data Mappings via Linear SVM. *Journal of Machine Learning Research*. 2010;11(48):1471-1490.

27. Tin Kam Ho. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Machine Intell*. 1998;20(8):832-844. doi:10.1109/34.709601

28. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM; 2016:785-794. doi:10.1145/2939672.2939785

29. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning: With Applications in R*. Springer US; 2021. doi:10.1007/978-1-0716-1418-1

30. Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. In: Guyon I, Luxburg UV, Bengio S, et al., eds. *Advances in Neural Information Processing Systems*. Vol 30. Curran Associates, Inc.; 2017. https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

31. Apley DW, Zhu J. Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2020;82(4):1059-1086. doi:10.1111/rssb.12377

32. Friedman JH, Popescu BE. Predictive learning via rule ensembles. *The Annals of Applied Statistics*. 2008;2(3):916-954. doi:10.1214/07-AOAS148

33. Jung RE, Haier RJ. The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*. 2007;30(2):135-154. doi:10.1017/S0140525X07001185

34. Santarnecchi E, Emmendorfer A, Pascual-Leone A. Dissecting the parieto-frontal correlates of fluid intelligence: A comprehensive ALE meta-analysis study. *Intelligence*. 2017;63:9-28. doi:10.1016/j.intell.2017.04.008

35. Wechsler D. Wechsler Adult Intelligence Scale--Fourth Edition. Published online November 12, 2012. doi:10.1037/t15169-000

36. Glasser MF, Smith SM, Marcus DS, et al. The Human Connectome Project's neuroimaging approach. *Nat Neurosci*. 2016;19(9):1175-1187. doi:10.1038/nn.4361

37. Fischl B, Salat DH, Busa E, et al. Whole Brain Segmentation. *Neuron*. 2002;33(3):341-355. doi:10.1016/S0896-6273(02)00569-X

38. Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci

using standard anatomical nomenclature. *NeuroImage*. 2010;53(1):1-15. doi:10.1016/j.neuroimage.2010.06.010

39. Fischl B. FreeSurfer. *NeuroImage*. 2012;62(2):774-781. doi:10.1016/j.neuroimage.2012.01.021

40. Poldrack RA, Huckins G, Varoquaux G. Establishment of Best Practices for Evidence for Prediction: A Review. *JAMA Psychiatry*. 2020;77(5):534-540. doi:10.1001/jamapsychiatry.2019.3671

41. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*. 1979;86(2):420-428. doi:10.1037/0033-2909.86.2.420

42. Ji JL, Spronk M, Kulkarni K, Repovš G, Anticevic A, Cole MW. Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage*. 2019;185:35-57. doi:10.1016/j.neuroimage.2018.10.006

43. Marcus D, Harwell J, Olsen T, et al. Informatics and Data Mining Tools and Strategies for the Human Connectome Project. *Frontiers in Neuroinformatics*. 2011;5:4. doi:10.3389/fninf.2011.00004

44. Scullin MK, Bliwise DL. Sleep, Cognition, and Normal Aging: Integrating a Half Century of Multidisciplinary Research. *Perspect Psychol Sci*. 2015;10(1):97-137. doi:10.1177/1745691614556680

Data Security Agreement

Provisional Paper Title: Multimodal brain-based predictive model for cognitive changes in middle-aged adults
Proposing Author: Narun Pat
Today's Date: 11/11/2021

☒	I am current on Human Subjects Training (CITI (www.citiprogram.org) or equivalent)
☒	My project is covered by the Duke ethics committee OR I have /will obtain ethical approval from my home institution.
☒	I will treat all data as "restricted" and store in a secure fashion. My computer or laptop is: a) encrypted (recommended programs are FileVault2 for Macs, and Bitlocker for Windows machines) b) password-protected c) configured to lock-out after 15 minutes of inactivity AND d) has an antivirus client installed as well as being patched regularly.
☒	I will not "sync" the data to a mobile device.
☒	In the event that my laptop with data on it is lost, stolen or hacked, I will immediately contact Moffitt or Caspi.
☒	I will not share the data with anyone, including my students or other collaborators not specifically listed on this concept paper.
☒	I will not post data online or submit the data file to a journal for them to post. *Some journals are now requesting the data file as part of the manuscript submission process. Study participants have not given informed consent for unrestricted open access, so we have a managed-access process. Speak to Temi or Avshalom for strategies for achieving compliance with data-sharing policies of journals.*
☒	I will delete all data files from my computer after the project is complete. Collaborators and trainees may not take a data file away from the office. This data remains the property of the Study and cannot be used for further analyses without an approved concept paper for new analyses.
☒	I have read the Data Use Guidelines and agree to follow the instructions.

Signature: Narun Pat

CONCEPT PAPER RESPONSE FORM

A

Provisional Paper Title	Multimodal brain-based predictive model for cognitive changes in middle-aged adults
Proposing Author	Narun Pat
Other Contributors	Tracy Melzer, Richie Poulton, Annchen Knodt
Potential Journals	Brain, Cerebral Cortex, Neuroimage, Cortex, Human Brain Mapping, Neurobiology of Aging, Journals of Gerontology
Today's Date: 11/7/2021	
Intended Submission Date: 12/31/2023	

Please keep one copy for your records and return one to the proposing author

B. To be completed by potential co-authors:

☒	Approved
☐	Not Approved
☐	Let's discuss, I have concerns

Comments: Click here to enter text

Please check your contribution(s) for authorship:

☒	Conceptualizing and designing the longitudinal cohort study
☒	Conceptualizing data collection protocols and creating variables
☒	Data collection
☐	Conceptualizing and designing this specific paper project
☐	Statistical analyses and interpretation (or reproducibility check)
☐	Writing
☒	Reviewing manuscript drafts
☒	Final approval before submission for publication
☒	Agreement to be accountable for the work
☐	Acknowledgment only, I will not be a co-author

Signature: Richie Poulton