Brain Structure and Function Predict Adherence to an Exercise Intervention in Older Adults

TIMOTHY P. MORRIS[1], AGNIESZKA BURZYNSKA[2], MICHELLE VOSS[3], JASON FANNING[4], ELIZABETH A. SALERNO[5], RUCHIKA PRAKASH[6], NEHA P. GOTHE[7,8], SUSAN WHITFIELD-GABRIELI[1,9], CHARLES H. HILLMAN[1,10], EDWARD MCAULEY[7,8], and ARTHUR F. KRAMER[1,7]

[1]Department of Psychology, Northeastern University, Boston, MA; [2]Department of Human Development and Family Studies, Colorado State University, Fort Collins, CO; [3]Deptartment of Psychology, University of Iowa, Iowa City, IA; [4]Department of Health and Exercise Science, Wake Forest University, Winston-Salem, NC; [5]Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine in St. Louis, St. Louis, MO; [6]Department of Psychology, Ohio State University, Columbus, OH; [7]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana–Champaign, Urbana, IL; [8]Department of Kinesiology and Community Health, University of Illinois at Urbana–Champaign, Urbana, IL; [9]McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA; and [10]Department of Physical Therapy, Movement, and Rehabilitation Sciences, Northeastern University, Boston, MA

BASIC SCIENCES

ABSTRACT

MORRIS, T. P., A. BURZYNSKA, M. VOSS, J. FANNING, E. A. SALERNO, R. PRAKASH, N. P. GOTHE, S. WHITFIELD-GABRIELI, C. H. HILLMAN, E. MCAULEY, and A. F. KRAMER. Brain Structure and Function Predict Adherence to an Exercise Intervention in Older Adults. *Med. Sci. Sports Exerc.*, Vol. 54, No. 9, pp. 1483–1492, 2022. **Introduction:** Individual differences in brain structure and function in older adults are potential proxies of brain reserve or maintenance and may provide mechanistic predictions of adherence to exercise. We hypothesized that multimodal neuroimaging features would predict adherence to a 6-month randomized controlled trial of exercise in 131 older adults (age, 65.79 ± 4.65 yr, 63% female), alone and in combination with psychosocial, cognitive, and health measures. **Methods:** Regularized elastic net regression within a nested cross-validation framework was applied to predict adherence to the intervention in three separate models (brain structure and function only; psychosocial, health, and demographic data only; and a multimodal model). **Results:** Higher cortical thickness in somatosensory and inferior frontal regions and less surface area in primary visual and inferior frontal regions predicted adherence. Higher nodal functional connectivity (degree count) in default, frontoparietal, and attentional networks and less nodal strength in primary visual and temporoparietal networks predicted exercise adherence ($r = 0.24$, $P = 0.004$). Survey and clinical measures of gait and walking self-efficacy, biological sex, and perceived stress also predicted adherence ($r = 0.17$, $P = 0.056$); however, this prediction was not significant when tested against a null test statistic. A combined multimodal model achieved the highest predictive strength ($r = 0.28$, $P = 0.001$). **Conclusions:** Our results suggest that there is a substantial utility of using brain-based measures in future research into precision and individualized exercise interventions older adults. **Key Words:** AGING, FUNCTIONAL CONNECTIVITY, BRAIN RESERVE, PREDICTION, MACHINE LEARNING, AEROBIC EXERCISE

Physical activity and structured exercise have received a lot of attention as potential efficacious interventions to improve or maintain brain health with advancing age (1). Despite millions of government and private dollars being spent on understanding how physical activity can improve or maintain brain function across the life span, over a third of the U.S. population do not engage in sufficient physical activity (2), a statistic that continues to increase (2). This physical inactivity pandemic (3) is estimated to cost private and public health care systems $53.8 billion per year and a further $13.7 billion in productivity losses due to physical inactivity-related deaths each year (4).

Previous research on understanding the adoption and maintenance of physical activity has leveraged psychological and psychosocial theories (5). In experimental studies, several psychosocial, behavioral, and demographic measures were shown to correlate with exercise adherence, such as self-efficacy (6), self-regulation (7), social support, perceived benefits and biological sex (males adhered more than females) (8), higher baseline physical activity outside the intervention (9), and depression, fatigue, and general perceived health (10). Additionally,

Address for correspondence: Timothy P. Morris, Ph.D., Department of Psychology, Northeastern University, 606 ISEC, 360 Huntington Avenue, Boston, MA 02115; E-mail: t.morris@northeastern.edu.
Submitted for publication August 2021.
Accepted for publication April 2022.
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/22/5409-1483/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000002949

greater cognitive resources, particularly executive functions, known to depend on cortical structural integrity (11), have been associated with engagement in exercise (12,13). Although theory-based programs are generally successful, optimizable, and adaptable, precision medicine approaches will improve the cost efficiency of interventions and health outcomes for individuals who show poor adherence to one-size-fits-all approaches. That is, predicting from a set of variables, who is more or less likely to adhere to an intervention from the outset, will allow the practitioner to provide alternative and individualized interventions before individuals having to either demonstrate poor adherence or present with deteriorating health.

Within the intervention setting, several recent studies have demonstrated that preintervention brain structure, known to be critical in supporting executive functions (11), is associated with individual differences in adherence to exercise interventions. Specifically, regions of the prefrontal, temporal, and somatosensory cortices have been associated with adherence to each respective exercise intervention (14,15). Relatedly, several reviews and observational studies have suggested that the relationship between exercise and the brain is bidirectional (13,16). This hypothesis leverages the concepts of cognitive reserve and brain maintenance (17), such that there is a circular nature between greater cognitive and brain resources and higher participation in complex exercise behaviors, which in turn helps with the maintenance and upkeep of cognitive and brain health. Although rarely measured directly, proxy measures of brain reserve and brain maintenance include brain structure and brain functional connectivity (for a review, see Stern et al. [17]). In terms of function, numerous applications of functional and anatomical connectivity have led to the observation that the brain is organized into large-scale functional networks (18). A number of these networks, such as the frontoparietal control network (FPCN), the default mode network (DMN), and the dorsal attention network (DAN), change with age (19) and are thought to be particularly important in age-related cognitive decline (20). These networks also subserve internally and externally directed cognition (21), including broadly defined executive function. Given the heterogeneity of these brain networks and their importance in higher-order cognition, individual patterns of functional connectivity in older adults may be predictive of exercise adherence. Indeed, prior studies on mindfulness and meditation have demonstrated that functional connectivity within the DMN and in frontal and temporal nodes are strong predictors of adherence (22).

The term "prediction" is used in several ways in the literature and can commonly refer to either the correlation of one variable in a group at one point in time with another variable in that same group at another point in time (within-sample correlation), or it can refer to a generalizable model that makes predictions on out-of-sample participants (23). Prior studies on psychosocial and behavioral "predictions" of exercise adherence have typically used correlation and therefore their results are likely overly optimistic (24). This overestimation and poor generalizability is compounded further when one considers they only explain relatively small amount of variance (<20%) in adherence (25,26). Consequently, calls have been made to introduce prediction statistics into psychological and cognitive neuroscience research to improve the generalizability of the results, specifically when using brain-based metrics (23). In the case of exercise interventions, for example, being able to successfully predict adherence to exercise interventions with generalizable results before engagement in the intervention would allow for the optimization or individualization of an alternative intervention for those who are predicted to poorly adhere. For example, providing those who are predicted to adhere poorly with a health coach (27) or just-in-time messaging (28) paves the way toward efficacious precision medicine approaches in health-based settings, when time can rarely be lost, and cost efficiency is of upmost importance. This is also important as engagement in effective exercise interventions may lead to the successful maintenance of physical activity after cessation of the intervention (29), and so predicting adherence to the exercise intervention may also provide insights into sustained exercise behaviors. Moreover, if brain-based metrics are shown to predict adherence, they can then provide mechanistic understanding of this complex behavior. Under the National Institutes of Health stage model of intervention development, such information could then be leveraged to design novel interventions that improve exercise adherence by targeting certain predictive mechanisms and testing whether their modulation improves adherence.

In a secondary analysis of data from a randomized control trial of exercise in older adults, the objectives of this study were to predict adherence to a 6-month exercise intervention and an active control in a structured and supervised group-based intervention. We hypothesized that multiple metrics of brain structure and function in FPCN, DMN, and DAN and regions (prefrontal, temporal, and parietal) would predict adherence to the exercise intervention and that these measures would augment traditional behavioral and psychosocial measures. To test this, we used a whole-brain, data-driven machine learning approach.

METHODS

Participants

This study is a secondary analysis of data from participants who participated in a 6-month randomized controlled exercise trial (clinical study identifier: NCT01472744, November 16, 2011). The study procedures were approved by the University of Illinois Institutional Review Board, and written informed consent was obtained from all participants before any research activities. Healthy but low-active older adults were recruited in Champaign County, Illinois. Two hundred and forty-seven (169 women) low-active older adults met the inclusion criteria for the initial clinical trial, of which 165 underwent structural and functional magnetic resonance imaging (MRI) at baseline, before any involvement in intervention sessions, and 131 had complete data across all variables and were included in this analysis (see further exclusion criteria based on adherence below). Participants in the initial trial were randomized to one of four intervention conditions: a walking intervention, the same walking intervention plus a dietary supplement designed to enhance lean muscle mass, a dancing intervention, and an active control consisting of a stretching and toning intervention. All intervention groups met for approximately 1 h, three times per week for 6 months, and for the purpose of this analysis, we analyzed our main outcome (adherence) with all four conditions combined. A one-way ANOVA revealed no significant differences in adherence between the conditions ($F_{1,129}$ = 1.11, P = 0.294), and no significant differences in baseline characteristics between conditions existed (see Table S1, Supplemental Digital Content 1, Demographics stratified by group assignment, http://links.lww.com/MSS/C597). Further, we replicate the pattern results in the most common intervention condition (walking) only as a sensitivity analysis with a reduced sample size (n = 59) to ensure our results were not affected by condition assignment (see Table S2, Supplemental Digital Content 2, Prediction models in the walking and walking + group only, http://links.lww.com/MSS/C598). Initial inclusion criteria included being between the ages of 60 and 80 yr old; free from psychiatric and neurological illness, including no history of stroke, transient ischemic attack, or head trauma; scored >23 on the Mini-Mental State Exam, >21 on a Telephone Interview of Cognitive Status questionnaire, and <10 on the Geriatric Depression Scale; at least 75% right-handed based on the Edinburgh Handedness Questionnaire (a criterion related to functional magnetic resonance imaging (MRI) analyses); demonstrated normal or corrected-to-normal vision of at least 20/40 and no color blindness; screened for safe participation in an MRI environment (e.g., no metallic implants that could interfere with the magnetic field or cause injury and no claustrophobia); and reported to have participated in no more than two bouts of moderate exercise per week within the past 6 months (with the goal of recruiting low-active older adults). Table 1 contains complete characterization of the study participants included in this analysis. All methods were carried out in accordance with the Declaration of Helsinki.

Adherence

Our primary outcome, adherence, was modeled as percentage attendance to the weekly intervention sessions. These supervised sessions were scheduled three times per week and lasted approximately for 1 h. The number and the frequency of the sessions were consistent across all four conditions. For this analysis, participants (n = 34) who dropped out of the study (i.e., did not complete the intervention) were excluded, as were those who failed to attend at least half of the intervention sessions

TABLE 1. Participant characteristics.

N	131
Age, mean ± SD	65.79 ± 4.65
Gender: female (%)	63 (71.6)
Race (%)	
White	79 (89.7)
African American or Black	7 (8)
Asian	2 (2.3)
Years of education, mean ± SD	15.89 ± 2.66
Adherence (%)	81 (12)

(n = 12) as our aim was to capture and predict variations in adherence to the entire 6-month intervention.

Psychosocial, Physical Function and Activity, Health, and Cognitive Features

A comprehensive battery of psychosocial, physical, and cognitive assessments was completed by each participant preintervention. A complete table of these assessments is found in Table S3 (see Supplemental Digital Content 3, Complete list of behavioral variables, http://links.lww.com/MSS/C599). Eighty-four assessments were initially included in this analysis. In brief, these assessments included self-reported psychosocial questionnaires gauging participants self-efficacy (global, exercise, and gait self-efficacy), leisure time activity, perceived sleep quality, anxiety, depression and self-worth/esteem, barriers to exercise, self-regulation, stress, loneliness, and subjective memory. All psychosocial measures were taken at week 1 or in some instances repeated at week 3. A set of physical function tests were also collected at week 1, which included a stair climb test, arm curl, sit and reach, and back scratch. Seven days of accelerometry capturing objective measures of time spent sedentary, time spent in light or moderate to vigorous aerobic physical activity, and average daily step counts were also collected. The procedures to capture, preprocess, and validate these measures can be found in a previous publication (30). A measure of cardiorespiratory fitness from a complete cardiopulmonary exercise test was also included as were measures of body composition (see our previous work [31] for a detailed description of the methodology of these measures). Finally, a battery of neuropsychological tasks were completed in week 1, which included numerous assessments of vocabulary, abstract, inductive and visuospatial reasoning, memory, and perceptual speed, taken from the Virginia Cognitive Aging Project (32). Intervention condition assignment was also included as a feature.

Neuroimaging Features

Magnetic resonance imaging: acquisition. Participants undertook an MRI scanning session in a 3T Siemens Trio Tim system with a 12-channel head coil before the intervention. High-resolution structural MRI scans were acquired using 3D MPRAGE T1-weighted sequences (repetition time = 1900 ms, echo time = 2.32 ms, inversion time = 900 ms, flip angle = 9°, matrix = 256 × 256, field of view = 230 mm, 192 slices, resolution = 0.9 × 0.9 × 0.9 mm; GRAPPA acceleration factor 2). T2*-weighted resting state echo-planar imaging data were obtained with the following parameters: 6 min, repetition time = 2 s, echo time = 25 ms, flip angle = 80°, 3.4 × 3.4 mm^2 in-plane resolution, 35 4-mm-thick slices acquired in ascending order, Grappa acceleration factor = 2, 64 × 64 matrix). Structural and resting state functional images were acquired with these scanning parameters, and the preprocessing and analyses of each respective modality are outlined in the following two sections.

BASIC SCIENCES

Structural MRI preprocessing and analyses. Cortical reconstruction and image segmentation and estimation of the cortical surface models were performed using the freely available FreeSurfer software v.5.3 (http://surfer-nmr.mgh.harvard.edu/). For the preprocessing of the cortex, a three-dimensional surface model was created using the "recon-all" surface-based stream. Automated Talairach transformation and intensity normalization were followed by nonbrain tissue removal, tessellation of the gray and white matter boundary, and automated topology correction. Finally, surface deformation enabled the detection of tissue boundaries: gray–white and gray–CSF borders. The cortical surfaces were then inflated and registered to a spherical atlas that used individual cortical folding patterns to match cortical geometry across participants. Cortical thickness was calculated at each vertex in the cortex as a measure of the distance between the white and the pial surfaces, and cortical surface area was calculated by averaging the area of all faces that meet at a given vertex on the white matter surface. We chose to analyze cortical thickness and surface area separately given their genetic independence and sensitivity to clinical and aging outcomes (33). Automatic labeling per the Desikan–Killiany cortical parcellation scheme was performed, and average cortical thickness and cortical surface area were calculated within each parcellation, resulting in 136 structural features (68 features per modality) to be used for feature selection.

Resting state functional connectivity preprocessing and analyses. Preprocessing of the functional resting state data was performed using the CONN-toolbox v.19c (34), relying upon SPM v.12 (Wellcome Department of Imaging Neuroscience, UCL, London, UK) in MATLAB R2019a (The MathWorks Inc., Natick, MA). The default preprocessing pipeline implemented in CONN was performed, which consists of the following steps: functional realignment and unwarping, slice timing correction, outlier identification, segmentation (into gray matter, white matter, and cerebrospinal fluid tissue), and normalization into the standard Montreal Neurologic Institute space with 2-mm isotropic voxels for functional data and 1 mm for anatomical data, using fourth-order spline interpolation. Finally, functional scans were spatially smoothed using a 6-mm Gaussian kernel. During the outlier detection step, acquisitions with framewise displacement above 0.9 mm (per several prior publications in studies with older adults who are more susceptible to movement within the scanner) (35–37) or global BOLD signal changes above 5 SD were flagged as potential outliers using the Artifact Detection Tools (www.nitrc.org/projects/artifact_detect). Two participants were removed from the final analyses for having >40 volumes flagged. This cutoff was determined based on preserving at least 5 min of scanning time (38). Additionally, mean framewise displacement was calculated via the Jenkinson method and regressed out of the final analysis (see Statistical analysis section). This was done to be overconservative given that previous studies have shown a high degree of motion–behavior correlations (39), despite the fact that no motion variable was significantly correlated with adherence outcomes in our study (all $P > 0.1$). Denoising of the functional data was performed using a component-based correction method, CompCor (40), and temporal band-pass filtering (0.01–0.1 Hz) to remove physiological, subject-motion, and outlier artifacts. Linear regression was used to remove the effects of these artifacts on the BOLD time series for each voxel and each subject considering noise components from voxels within white matter and cerebrospinal fluid, estimated subject-motion parameters (three rotation and three translation parameters and six other parameters representing their first-order time derivatives), scrubbing, and constant and first-order linear session effects.

To prepare the functional data for feature selection, we parcellated the functional scans into a medium resolution 300-region (node) atlas with a 17-network parcellation scheme (41). A medium resolution parcellation was chosen to strike a balance between having sufficient dimensionality to capture predictive information within the functional connectivity matrix and having too many dimensions that would create a very large search space for the subsequent model building described below. Mean BOLD activity was calculated within each node and transformed into a 300 × 300 correlation matrix where the time series at each node was correlated with that of every other node. This matrix then underwent Fisher's Z transformation. Functional brain connectivity was summarized using degree count, a graph theory metric that represents whole-brain connection density for each node, per a previous publication (42). This metric reduces the dimensions of the functional connectivity data into a sparse matrix that better represents real-world graphs (43). Degree count for each participant was calculated at several thresholds using a fixed network cost (keeping the strongest 15%, 20%, and 25% of connections), and a final threshold was chosen through cross-validation (see Statistical analysis section). Degree count at each node here represents a single measure of the sum of connections between that node and every other node in the cortex characterizing its degree of connectedness within the cortical gray matter.

Power Analysis

We performed a power analysis (in R using the "pwr" package) on our sample size to ensure sufficient power was gained to detect a true effect. Based on our sample size (n = 131) and an assumed type I error rate of 0.05, we calculated an estimated 96% power to detect an effect size of 0.22 from a general linear model with 38 covariates (multimodal model). In the reduced sample size of n = 59 for the sensitivity analysis (see Table S2, Supplemental Digital Content 2, Prediction models in the walking and walking + group only, http://links.lww.com/MSS/C598), this power fell to 60% for the same effect size.

Statistical Analysis

To predict adherence to the exercise interventions, we used elastic net regression, a regularized (penalized) regression method within a nested cross-validation procedure. Elastic net is a data-driven machine learning regression method that applies a penalty term to each variable (feature) coefficient in the model, resulting in the shrinkage of that coefficient's value. This penalty term is a function of the overall weight of a given feature in a model and serves to improve the generalizability of the model results when applied to new data. Elastic net aims to avoid overfitting (a problem in statistics that leads to poor generalizability of study results) by producing less complex models by applying both L1 and L2 norm penalties to calculate the coefficients of each feature. Elastic net is a linear combination of both Ridge regression (L2 norm) and least absolute selection and shrinkage operate (Lasso) regression. The L1 penalty shrinks coefficients toward zero (Ridge). The L2 penalty, in the case of features that have no predictive value, shrinks their coefficients to exactly zero (Lasso), relative to the maximum likelihood estimates, effectivity removing them from the model. The predictors with nonzero coefficients are therefore interpreted as those that contained predictive information and contributed to the final model of predicting the outcome. This model is particularly useful in cases where correlated features are present (i.e., neuroimaging features and some psychosocial measures) as the combination of both L1 and L2 penalties will maintain groups of correlated features in the model (whereas the L1 penalty [Lasso] alone would remove all but one of the correlated features). The amount of shrinkage is determined via tuning of two hyperparameters λ_1 and λ_2, whereby via a grid search approach the optimal combination of penalties is "tuned" by running the models over and over and assessing the predictive performance at each level of the penalty. Results of the optimal hyperparameters for each model are found in Table 3. In this analysis, we used a nested cross-validation procedure where the data set is split into a 10-fold outer loop (to evaluate model performance) and a 10-fold inner loop (to tune the hyperparameters using grid search within each inner fold). Nested cross-validation avoids optimization bias that simple cross-validation could potentially suffer from when using the same folds to both tune the hyperparameters and test the prediction performance. In this case, the folds were kept consistent across each model by setting the same random seed, and all variables were centered and scaled within each inner loop and applied to the outer folds. Cross-validation is an important step in prediction modeling whereby the predictive performance of a model is assessed on left-out data (a model is trained in $n - 1$ folds and tested on the left-out fold in an iterative fashion), resulting in an unbiased measure of the prediction performance on unseen data. This step overcomes the issues of simple multiple linear regression whereby a good model (fit to the whole data set) may not be a good predictive model when applied to new data, resulting in generalizability issues.

First, to select the optimal degree threshold of the functional connectivity metric to use in the final models, functional connectivity–only models were trained over each threshold density separately, and the model that predicted the left-out outer folds with the smallest root-mean-squared error (RMSE) was selected to be used in the final models (for full results from this analysis, see Table S4, Supplemental Digital Content 4, Optimal degree count threshold, http://links.lww.com/MSS/C600). To reduce the features used in the final models, a selection by filtering approach was taken within the same cross-validation folds used in the final models to keep test and train data separate and to reduce overfitting. Here, features that did not correlate with the outcome at $P < 0.05$ within each fold were removed from the final feature set. Age, biological sex, and mean framewise displacement (functional connectivity only) were regressed out of the features. Three separate models were trained: 1) imaging features only; 2) psychosocial, physical function and activity, health, and cognitive features only; and 3) multimodal features containing all features. In the case of model 2, the combination of nested cross-validation and elastic net failed to predict adherence in the left-out outer folds, potentially because of the small number of features in the model ($P = 6$), and as such, for the nonimaging model only, we ran a simple 10-fold cross-validation model using Ridge regression (L1 norm penalty only). We provide several interpretable measures of model performance (prediction on the left-out folds) based on the observed versus predicted values: Pearson's r correlation; squared correlation; R^2; RMSE, which measures the average prediction error as the average difference between the observed and the predicted values; and the mean absolute error (MAE) as the average absolute difference between the observed and the predicted values. RMSE and MAE are related with MAE being less sensitive to outliers, and the lower the value, the better the model performance. To assess the significance of the prediction performance, 10,000 nonparametric permutations were performed on the correlation coefficient between the predicted and the observed values, consistent with the fMRI literature (44). Here, the outcome measure is randomly assigned to different subjects, and by using this new label assignment 10,000 times, we estimated the distribution of the test statistic. The P value of the permutation tests was then calculated as the proportion of sampled permutations that are greater or equal to the true prediction correlation (see Fig. S1, Supplemental Digital Content 5, Significance testing of the prediction performance, http://links.lww.com/MSS/C601). All statistics were performed in RStudio version 3.6.3 using "caret," "dplyr," "purr," "penalized," and "pensim" packages. Figures were generated using "ggplot2." Code used in this analysis can be found at https://github.com/tpmor-546/Adherence_pred.

RESULTS

Table 1 presents demographic details of the participants included in this analysis. Our sample consisted of mostly female, white, and educated participants, and the mean adherence was 81%.

Table 2 contains prediction metrics for each modality using elastic net models and ridge regression for the psychosocial/cognitive/health only model. The psychosocial and demographic model did not significantly predict adherence (r = 0.17, P = 0.056). The imaging (r = 0.24, P = 0.004) and multimodal models (r = 0.28, P = 0.001) did significantly predict adherence (Table 2 and Fig. 1).

Standardized coefficients for each feature selected in each model are found in Table 3. A summary figure of the multimodal model is found in Figure 2. For the functional connectivity features, higher nodal degree strength in lateral and

TABLE 2. Model performance metrics.

Model	r	R^2	RMSE	MAE	P
Multimodal	0.28	0.08	0.11	0.09	0.001
Imaging	0.24	0.06	0.11	0.09	0.004
Psych/cog/health	0.17	0.03	0.12	0.10	0.056

Performance metrics derived from nested cross-validation where the optimal hyperparameters for elastic net regression were tuned in an inner loop and used to predict adherence in the left-out outer loop. r and R^2 represent the Pearson's correlation and the squared correlation between the predicted and the observed values, respectively. RMSE represents the average difference between the observed and the predicted values (average prediction error), and MAE represents the absolute mean difference between the predicted and the observe values. The P value for each model is derived by comparing the correlation coefficient between the observed and the predictive values to a null distribution derived from 10,000 nonparametric permutations.

medial prefrontal regions, spanning both FPCN and DMN, predicted higher adherence to the intervention. Conversely, less nodal strength in visual, temporal, and somatomotor networks predicted higher adherence to the intervention. Higher cortical thickness in bilateral postcentral gyrus, inferior frontal gyrus of the left hemisphere, and the left frontal pole predicted higher adherence, whereas lower cortical thickness in the right cingulate and lower surface area in bilateral occipital gyrus and left inferior frontal gyrus predicted higher adherence. Although the psychosocial and demographic model alone did not significantly predict adherence, those features included in the multimodal model that was predictive of adherence (given their nonzero coefficient) included self-efficacy for walking, barrier-specific self-efficacy, strength self-esteem, biological sex, and employment status (Table 3).

DISCUSSION

In this study, we demonstrated that brain-based measures of functional connectivity, cortical thickness, and surface area predicted future adherence to a structured group-based exercise intervention in older adults. Using a machine learning framework that applies penalized regression with cross-validation, we also replicated earlier studies indicating that aspects of self-efficacy and biological sex are predictive of adherence to exercise and that, together, multimodal features provide the numerically strongest predictive value.

Currently, our model performance remains far from having utility in a clinical setting. Although we replicated the finding that traditional psychosocial and demographic measures predict adherence to exercise, albeit in a model that was not significant, the ability of neuroimaging features alone to predict adherence to a structured group-based exercise intervention suggests a substantial utility of these measures for future research into precision medicine and adaptive intervention approaches. Additionally, the numerical increase in model performance when combined suggests that multimodal features provide independent relevant information in the prediction of exercise adherence. Multimodal prediction models have been shown to outperform unimodal models in prior work also (42,45). It is possible that the numerical increase in prediction performance from multimodal features is due to individual features capturing distinct aspects of complex behaviors linked to adherence to exercise, which unimodal features alone may not capture. Additionally, the use of multimodal imaging metrics likely limits the effect of scanning or preprocessing artifacts from any given modality. Notwithstanding, although the correlations between predicted and observed values were higher in the multimodal model, the absolute difference between the predicted and the observed values (MAE) between the imaging model and the multimodal model was numerically negligible. Given that the psychological, health, and demographic model was not significant when tested against a null distribution of the test statistic, it is likely that the imaging features were providing most of the predictive signal.

It is important to highlight simple numerical comparisons between our results (i.e., R^2), and those of prior research on correlates of exercise adherence should not be compared directly because of the rigor of prediction modeling compared with simple within-sample inferential statistics. Our prediction methodology is implemented with the explicit goal of improving generalizability to new unseen data (23), whereas prior studies using within-sample correlational statistics are prone to biased estimates that may not generalize well to new data.

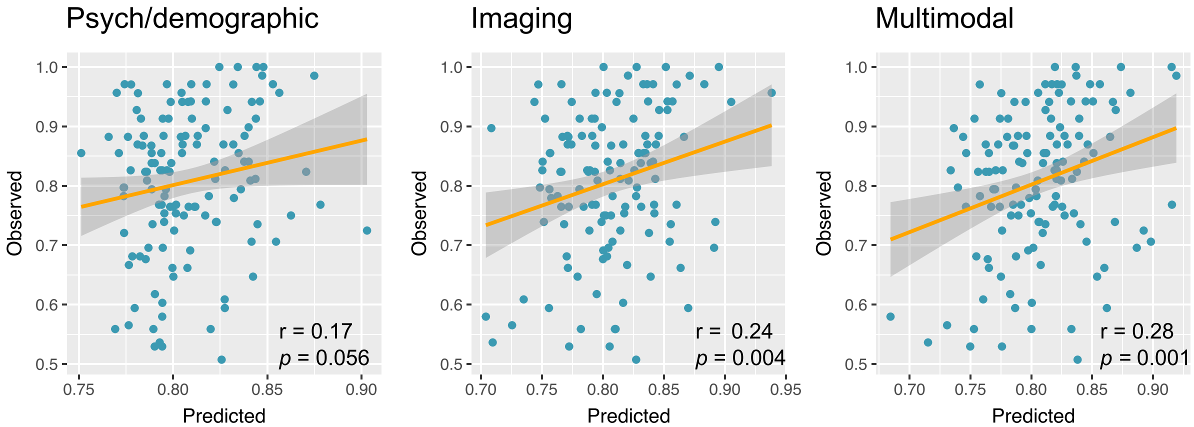

FIGURE 1—Prediction performance (*r*) of each model (*top row*). *Top row*: *r* represents the correlation between the predicted and the observed adherence values via elastic net regression with nested cross-validation. The multimodal model numerically performed better than the unimodal models alone.

TABLE 3. Data-driven features.

Feature	Psych/Cog/Health λ1	Psych/Cog/Health λ2	Imaging λ1	Imaging λ2	Multimodal λ1	Multimodal λ2
	N/A	98.5	0.1	100	0.46	56
lh_extrastriate_centralvisual			−0.00048		0.9136	
lh_lateralPFC_FPCNa			0.00259		0.00156	
lh_lateralPFC_DMNb			0.00572		0.00558	
lh_ventralPFC_DMNb			0.00553		0.00572	
lh_temporalparietal			−0.00687		−0.00765	
rh_extrastriatesuperior_perivisual			−0.00326		−0.00243	
rh_S2_SMb			−0.00531		−0.00594	
rh_parsopercularis_VANa			−0.00522		−0.00332	
rh_frontalmedial_VANa			0.00736		0.00632	
rh_medialprefrontal_DMNa			0.00635		0.00636	
rh_temporalparietal			−0.00131		0	
rh_temporalparietal			−0.00975		−0.0112	
lh_parsorbitalis_thickness			0.00512		0.00435	
lh_parstriangularis_thickness			0.00169		0.00049	
lh_postcentral_thickness			0.00627		0.00779	
rh_isthmuscingulate_thickness			−0.01194		−0.01149	
rh_postcentral_thickness			0.00453		0.00476	
rh_frontalpole_thickness			0.00153		0.00104	
lh_cuneus_area			−0.00218		0	
lh_parsorbitalis_area			−0.00464		−0.00798	
lh_pericalcarine_area			−0.00731		−0.00684	
rh_lateraloccipital_area			−0.00563		−0.04543	
Self-efficacy for walking (w3)	−0.00073				−0.01009	
Barriers self-efficacy (w1)	0.00014				0.00116	
Big Five; Conscientiousness	0.00068				0	
Avg time in light exercise (7 d)	0.87631				0	
Biological sex	−0.04877				−0.02005	
Strength self-esteem	−0.00059				−0.00049	
Employment status	0.00861				0.00561	

Data-driven features derived from selection by filtering approach within the cross-validation framework, which were used as initial input features in the elastic net models. Imaging feature names are constructed as follows: "hemishere_region_network/structural measure." Network is derived from the 17-network cortical parcellation from Yeo et al. (18) via automatic labeling using the Schaefer 300 atlas. Coefficients from each prediction model represent the standardized and penalized coefficient used to predict Adherence. A value of 0 means that the feature was effectivity removed from the model due to having little or no predictive value. λ1 and λ2 represent the optimal cross-validated hyperparameters for each model. λ1 is "N/A" for the psych/cog/health model as simple cross-validation, and ridge regression (L2 norm penalty only) was used to predict adherence as the elastic net with nested cross-validation failed to find a solution to predict adherence with these features only. Biological sex was coded as 2 for women and 1 for men (i.e., a negative coefficient means that men adhered better than women). Employment status was coded as 1–7, with 1 being full time employed, 2 part time, 3 retired, part time, 4 retired, 5 laid off/unemployed.

In a research setting, the application of prediction modeling could aid in the development of adaptive intervention strategies. For example, in those who can successfully engage in a theory-based intervention, positive results have been shown. However, successful engagement in long-term interventions is nontrivial and not all individuals will adhere. If one could predict these individuals from the outset, the redeployment of resources to tailor the intervention in these individuals with the goal of improving adherence and subsequent efficacy could be done. For example, our results could be leveraged in future prospective studies to test whether adaptive interventions based on preintervention characteristic and brain signatures (our predictions) result in better adherence and intervention efficacy. Concomitantly, our results provide several testable mechanistic pathways through which individual differences in brain structure and function can affect exercise adherence. For example, if default mode and executive control connectivity are important predictors of exercise adherence, then mindfulness meditation training could be used as a primer to enhance network connectivity before or during an exercise intervention, as this has been shown to modulate the connectivity of these networks (46). Future studies could test whether the modulation of these measures results in the enhancement of adherence to prospectively test these predictions. Additionally, successful engagement in an exercise intervention may also be related to the long-term maintenance of exercise behaviors beyond the intervention (29). Consequently, applying this type of prediction modeling to recent large population and cohort studies (e.g., UK Biobank or Human Connectome Project) with imaging and self-report physical activity data could help one determine the proportion and characteristics of individuals who would require additional tailoring when designing future exercise interventions.

Regarding our imaging analysis, we took a whole-brain approach to the prediction of exercise. Our results, however, could be leveraged to validate these results in future hypothesis-driven studies. For example, nodes with high global functional connectivity predictive of adherence in our study were found in the medial and lateral prefrontal cortex within the DMN and FPCN networks, two related networks that contribute to internally and externally directed cognition (21) and in the implementation of executive control processes to maintain goals and inhibit distractions, respectively (47). Similarly, cortical structure in postcentral and inferior temporal regions has been shown to be associated with executive functions, especially in healthy aging (11). As such, direct study of the interplay between these networks at rest or during executive function tasks that engage these networks may provide stronger predictive utility.

Numerous randomized controlled trials of exercise in aging have demonstrated that exercise can improve (with mixed effect sizes) cognitive function and positively affect brain structure and functional connectivity (48). More recently, the hypothesized bidirectional relationship between exercise and brain and cognition has been tested (13). In our current study, individual differences in proxy measures of cognitive and brain reserve (functional connectivity nodes, cortical thickness, and surface area, respectively) within primary information processing networks and regions were predictive of exercise adherence. Prior exercise interventions have shown intervention-mediated increases in the functional connectivity of somatosensory networks in older adults (49). Our results, therefore, potentially provide supportive evidence in favor of the hypothesis that the relationship between exercise and brain is bidirectional (50). Notwithstanding, our prediction models need to be tested in prospective and experimental studies to conclude a causal association. Furthermore, these predictions need to be validated in much larger and more generalizable samples.

Several limitations to this study mean that our results should be interpreted with reasonable caution. First, the complexity of adherence to a 6-month exercise intervention may not be completely captured in a single variable representing the percentage attendance to the intervention settings, and so additional explained variance not captured by our measures may be due to other reasons such as scheduling conflicts related to family, work, or breaks due to discomfort. Second, it is of note that inferential guarantees regarding variable estimates (coefficients) in penalized regression models cannot be made, i.e., traditional P values or confidence intervals for each

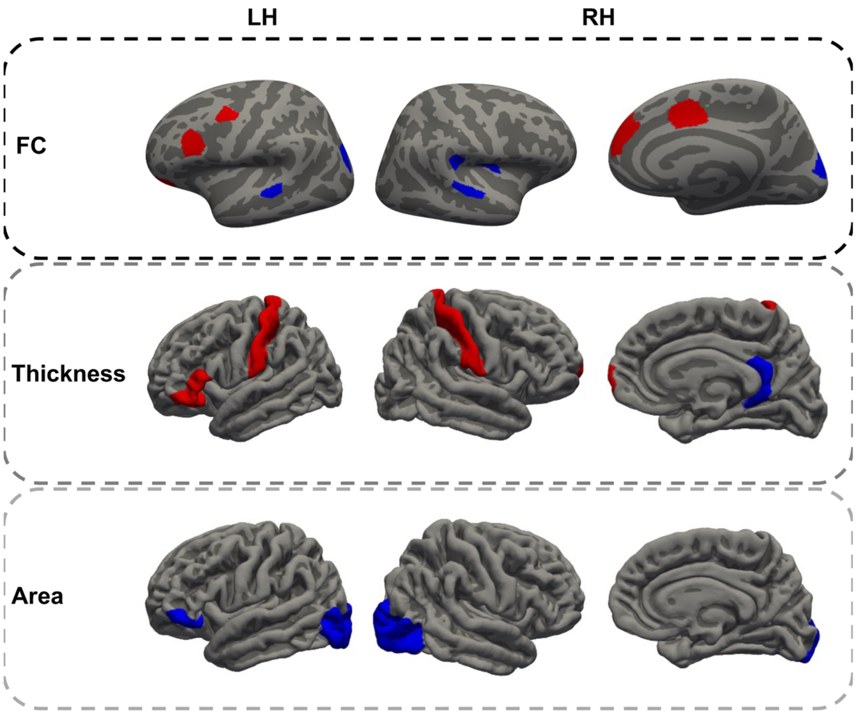

FIGURE 2—Summary figure depicting the data-driven features of brain structure and function in the multimodal model (best performing model). For the functional connectivity features (*top left*), each feature represents degree count (number of connections that each node has with every other node in the cortical parcellation). That is, *red features* (positive nodes) are interpreted as the higher the connectivity between that node and every other node in the cortex, the higher the adherence. *Blue features* (negative nodes) suggest that the less connections between these nodes and the rest of the brain, the higher the adherence. Mapping these regions to intrinsic resting state functional networks (17 Yeo networks via the Schaefer 300 atlas) reveals a broad pattern of positive nodes in the DMN and FPCN and negative nodes in primary information processing networks. For cortical brain structure, greater cortical thickness in bilateral postcentral gyrus and left inferior frontal gyrus and lower cortical surface area in bilateral occipital cortex and left inferior frontal gyrus were predictive of adherence.

estimate do not exist (51). Therefore, estimates produced by the elastic net model are biased and should not be interpreted as the population parameter. Although advances in postselection inference methods for the Lasso have been made (51,52), which allow for more valid confidence intervals and significance testing for Lasso estimates, such methods do not exist for elastic net, the use of which was important in our case given the collinearity of the predictors in our study. Third, our sample size was relatively small for machine learning. As such, testing on completely left-out subjects (test–train splits) was not feasible, and so we attempted to perform the most rigorous and generalizable approach possible via nested cross-validation. Fourth, we implemented a data-driven approach that used whole-brain imaging measures as features. This approach could be complimented or enhanced in future research using task-evoked MRI or network/region-specific measures to reduce the dimensionality of the input data. Fourth, given the difficulty and costs associated with this type of research and the vast number of measures that were available in this unique data set, the replication of this prediction in a completely independent validation data set is unavailable at this moment. Fifth, our participants were homogenous. Several factors have recently been highlighted as a reason why global physical inactivity rates continue to be low (3), with one being that most of the research on understanding physical activity behaviors has occurred in high-income countries, and so more work in diverse populations and low-income countries is needed to fully generalize this type of prediction. Lastly, our battery of neuropsychological tasks included executive function tests that measured constructs such as abstract, inductive, and visuospatial reasoning rather than other executive tasks like inhibitory control that may be more related to exercise adherence (12).

CONCLUSIONS

Our results showed that the combination of psychosocial, cognitive, and demographic and multimodal imaging metrics can predict adherence to an exercise intervention in older adults and provide independent relevant predictive value. Prospective testing of these predictions and their validation will allow researchers and eventually clinicians to leverage them through personalized medicine approaches.

The authors thank Anya Knecht, Susan Houseworth, Nancy Dodge, Hilly Tracy, Robert Weisshappel, and all of the Lifelong Brain and Cognition and Exercise Psychology Laboratory graduate students and staff for their help in participant recruitment and data collection.

This work was supported by the National Institute on Aging at the National Institutes of Health (R37 AG025667).

The authors declare no conflict of interest. The University of Illinois Institutional Review Board approved all procedures used in the study. All participants gave written informed consent before participation in any study procedures, all of which conformed to the Declaration of Helsinki for research involving human subjects. All authors agree to the contents of this manuscript and give consent for its publication.

The results of this study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute endorsement by the American College of Sports Medicine.

T. P. M. and A. F. carried out the conceptualization, design, analysis, and interpretation of data and wrote the manuscript. A. B., M. V., J. F., E. S., R. P., and N. G. performed data acquisition and substantial revision. S. W. G. analyzed and interpreted the data. C. H. designed the data and performed substantial revision. E. M. carried out the conceptualization, study design, and substantial revision.

All data will be provided upon reasonable request to the corresponding author, without reservation.

REFERENCES

1. Gomes-Osman J, Cabral DF, Morris TP, et al. Exercise for cognitive brain health in aging: a systematic review for an evaluation of dose. *Neurol Clin Pract*. 2018;8(3):257–65.
2. Du Y, Liu B, Sun Y, et al. Trends in adherence to the physical activity guidelines for Americans for aerobic activity and time spent on sedentary behavior among US adults, 2007 to 2016. *JAMA Netw Open*. 2019;2(7):e197597.
3. Pratt M, Ramirez Varela A, Salvo D, et al. Attacking the pandemic of physical inactivity: what is holding us back? *Br J Sports Med*. 2020; 54(13):760–2.
4. Ding D, Kolbe-Alexander T, Nguyen B, et al. The economic burden of physical inactivity: a systematic review and critical appraisal. *Br J Sports Med*. 2017;51(19):1392–409.
5. Brand R, Cheval B. Theories to explain exercise motivation and physical inactivity: ways of expanding our current theoretical perspective. *Front Psychol*. 2019;10:1147.
6. McAuley E, Mullen SP, Szabo AN, et al. Self-regulatory processes and exercise adherence in older adults: executive function and self-efficacy effects. *Am J Prev Med*. 2011;41(3):284–90.
7. de Bruin M, Sheeran P, Kok G, et al. Self-regulatory processes mediate the intention-behavior relation for adherence and exercise behaviors. *Health Psychol*. 2012;31(6):695–703.
8. Cadmus-Bertram L, Irwin M, Alfano C, et al. Predicting adherence of adults to a 12-month exercise intervention. *J Phys Act Health*. 2014; 11(7):1304–12.
9. Arkkukangas M, Söderlund A, Eriksson S, et al. One-year adherence to the Otago Exercise Program with or without motivational interviewing in community-dwelling older adults. *J Aging Phys Act*. 2018;26(3):390–5.
10. Flegal K, Kishiyama S, Zajdel D, et al. Adherence to yoga and exercise interventions in a 6-month clinical trial. *BMC Complement Altern Med*. 2007;7:37.
11. Burzynska AZ, Nagel IE, Preuschhof C, et al. Cortical thickness is linked to executive functioning in adulthood and aging. *Hum Brain Mapp*. 2011;33(7):1607–20.
12. Cheval B, Daou M, Cabral DAR, et al. Higher inhibitory control is required to escape the innate attraction to effort minimization. *Psychol Sport Exerc*. 2020;51:101781.
13. Cheval B, Orsholits D, Sieber S, et al. Relationship between decline in cognitive resources and physical activity. *Health Psychol*. 2020; 39(6):519–28.
14. Best JR, Chiu BK, Hall PA, et al. Larger lateral prefrontal cortex volume predicts better exercise adherence among older women: evidence from two exercise training studies. *J Gerontol A Biol Sci Med Sci*. 2017;72(6):804–10.
15. Gujral S, McAuley E, Oberlin LE, et al. Role of brain structure in predicting adherence to a physical activity regimen. *Psychosom Med*. 2018;80(1):69–77.
16. Audiffren M, André N. The exercise–cognition relationship: a virtuous circle. *J Sport Health Sci*. 2019;8(4):339–47.
17. Stern Y, Barnes CA, Grady C, et al. Brain reserve, cognitive reserve, compensation, and maintenance: operationalization, validity, and mechanisms of cognitive resilience. *Neurobiol Aging*. 2019;83:124–9.
18. Yeo BT, Krienen FM, Sepulcre J, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 2011;106(3):1125–65.
19. Grady C, Sarraf S, Saverino C, et al. Age differences in the functional interactions among the default, frontoparietal control, and dorsal attention networks. *Neurobiol Aging*. 2016;41:159–72.
20. Marek S, Dosenbach NUF. The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. *Dialogues Clin Neurosci*. 2018;20(2):133–40.
21. Spreng RN, Sepulcre J, Turner GR, et al. Intrinsic architecture underlying the relations among the default, dorsal attention, and frontoparietal control networks of the human brain. *J Cogn Neurosci*. 2013;25(1):74–86.
22. Saghayi M, Greenberg J, O'Grady C, et al. Brain network topology predicts participant adherence to mental training programs. *Netw Neurosci*. 2020;4(3):528–55.
23. Gabrieli JDE, Ghosh SS, Whitfield-Gabrieli S. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron*. 2015;85(1):11–26.
24. Yarkoni T, Westfall J. Choosing prediction over explanation in psychology: lessons from machine learning. *Perspect Psychol Sci*. 2017; 12(6):1100–22.
25. Rejeski WJ, Miller ME, King AC, et al. Predictors of adherence to physical activity in the lifestyle interventions and Independence for elders pilot study (LIFE-P). *Clin Interv Aging*. 2007;2(3):485–94.
26. Rhodes RE, Martin AD, Taunton JE. Temporal relationships of self-efficacy and social support as predictors of adherence in a 6-month strength-training program for older women. *Percept Mot Skills*. 2001;93(3):693–703.
27. Olsen JM, Nesbitt BJ. Health coaching to improve healthy lifestyle behaviors: an integrative review. *Am J Health Promot*. 2010;25(1): e1–12.
28. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med*. 2018;52(6):446–62.
29. Madigan CD, Fong M, Howick J, et al. Effectiveness of interventions to maintain physical activity behavior (device-measured): systematic review and meta-analysis of randomized controlled trials. *Obes Rev*. 2021;22(10):e13304.
30. Fanning J, Porter G, Awick EA, et al. Replacing sedentary time with sleep, light, or moderate-to-vigorous physical activity: effects on self-regulation and executive functioning. *J Behav Med*. 2017;40(2):332–42.
31. Voss MW, Weng TB, Burzynska AZ, et al. Fitness, but not physical activity, is related to functional integrity of brain networks associated with aging. *Neuroimage*. 2016;131:113–25.
32. Salthouse TA, Ferrer-Caja E. What needs to be explained to account for age-related effects on multiple cognitive variables? *Psychol Aging*. 2003;18(1):91–110.
33. Storsve AB, Fjell AM, Tamnes CK, et al. Differential longitudinal changes in cortical thickness, surface area and volume across the adult life span: regions of accelerating and decelerating change. *J Neurosci*. 2014;34(25):8488–98.
34. Whitfield-Gabrieli S, Nieto-Castanon A. Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect*. 2012;2(3):125–41.
35. Fitzhugh MC, Hemesath A, Schaefer SY, et al. Functional connectivity of Heschl's gyrus associated with age-related hearing loss: a resting-state fMRI study. *Front Psychol*. 2019;10:2485.
36. Pistono A, Guerrier L, Péran P, et al. Increased functional connectivity supports language performance in healthy aging despite gray matter loss. *Neurobiol Aging*. 2021;98:52–62.
37. Vieira BH, Rondinoni C, Garrido Salmon CE. Evidence of regional associations between age-related inter-individual differences in

resting-state functional connectivity and cortical thinning revealed through a multi-level analysis. *Neuroimage*. 2020;211:116662.
38. Van Dijk KR, Hedden T, Venkataraman A, et al. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J Neurophysiol*. 2010;103(1):297–321.
39. Siegel JS, Mitra A, Laumann TO, et al. Data quality influences observed links between functional connectivity and behavior. *Cereb Cortex*. 2017;27(9):4492–502.
40. Behzadi Y, Restom K, Liau J, et al. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage*. 2007;37(1):90–101.
41. Schaefer A, Kong R, Gordon EM, et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb Cortex*. 2018;28(9):3095–114.
42. Vogel JW, Vachon-Presseau E, Pichet Binette A, et al. Brain properties predict proximity to symptom onset in sporadic Alzheimer's disease. *Brain*. 2018;141(6):1871–83.
43. Power JD, Schlaggar BL, Lessov-Schlaggar CN, et al. Evidence for hubs in human functional brain networks. *Neuron*. 2013;79(4):798–813.
44. Shen X, Finn ES, Scheinost D, et al. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat Protoc*. 2017;12(3):506–18.
45. Mill RD, Winfield EC, Cole MW, et al. Structural MRI and functional connectivity features predict current clinical status and persistence behavior in prescription opioid users. *Neuroimage Clin*. 2021;30:102663.
46. Bauer CCC, Rozenkrantz L, Caballero C, et al. Mindfulness training preserves sustained attention and resting state anticorrelation between default-mode network and dorsolateral prefrontal cortex: a randomized controlled trial. *Hum Brain Mapp*. 2020;41(18):5356–69.
47. Spreng RN, Stevens WD, Chamberlain JP, et al. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage*. 2010;53(1):303–17.
48. Erickson KI, Hillman C, Stillman CM, et al. Physical activity, cognition, and brain outcomes: a review of the 2018 physical activity guidelines. *Med Sci Sports Exerc*. 2019;51(6):1242–51.
49. McGregor KM, Crosson B, Krishnamurthy LC, et al. Effects of a 12-week aerobic spin intervention on resting state networks in previously sedentary older adults. *Front Psychol*. 2018;9:2376.
50. Loprinzi PD, Herod SM, Cardinal BJ, et al. Physical activity and the brain: a review of this dynamic, bi-directional relationship. *Brain Res*. 2013;1539:95–104.
51. Lockhart R, Taylor J, Tibshirani RJ, et al. A significance test for the lasso. *Ann Stat*. 2014;42:413–68.
52. Lee JD, Sun DL, Sun Y, et al. Exact post-selection inference, with application to the lasso. *Ann Stat*. 2016;44:907–27.