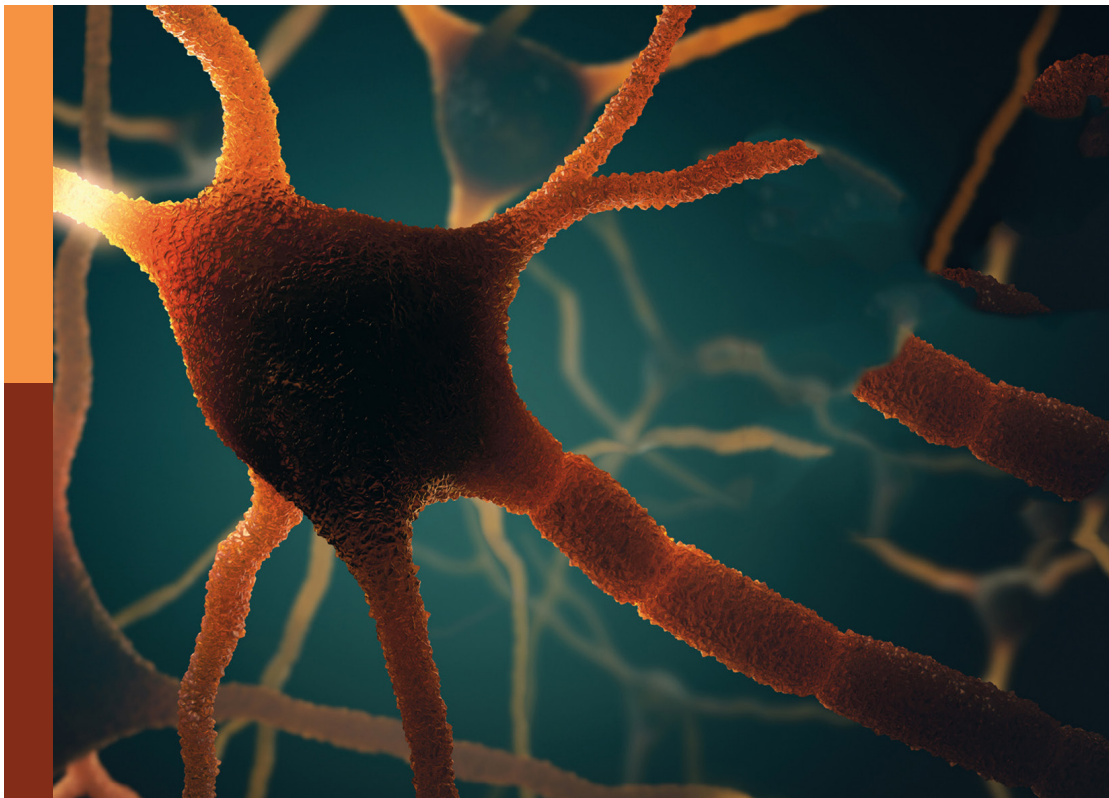
frontiers | Research Topics

Methods and applications in aging neuroscience

Edited by
Yang Jiang, Wai Haung Yu and S. Abid Hussaini

Published in
Frontiers in Aging Neuroscience



ISSN 1664-8714
ISBN 978-2-8325-2807-5
DOI 10.3389/978-2-8325-2807-5

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Methods and applications in aging neuroscience

Topic editors

Yang Jiang — University of Kentucky, United States

Wai Haung Yu — University of Toronto, Canada

S. Abid Hussaini — Columbia University Irving Medical Center, United States

Citation

Jiang, Y., Yu, W. H., Hussaini, S. A., eds. (2023). *Methods and applications in aging neuroscience*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2807-5

Table of contents

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 04 February 2022
doi: 10.3389/fnagi.2022.793298

Alteration of Neural Network Activity With Aging Focusing on Temporal Complexity and Functional Connectivity Within Electroencephalography

Momo Ando[1], *Sou Nobukawa*[1,2,3]*, *Mitsuru Kikuchi*[4,5] *and Tetsuya Takahashi*[5,6,7]

[1] Graduate School of Information and Computer Science, Chiba Institute of Technology, Narashino, Japan, [2] Department of Computer Science, Chiba Institute of Technology, Narashino, Japan, [3] Department of Preventive Intervention for Psychiatric Disorders, National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan, [4] Department of Psychiatry and Behavioral Science, Kanazawa University, Ishikawa, Japan, [5] Research Center for Child Mental Development, Kanazawa University, Kanazawa, Japan, [6] Department of Neuropsychiatry, University of Fukui, Yoshida, Japan, [7] Uozu Shinkei Sanatorium, Uozu, Japan

OPEN ACCESS

Edited by:
Yang Jiang,
University of Kentucky, United States

Reviewed by:
Cornelis Jan Stam,
VU Medical Center, Netherlands
Fatemeh Hadaeghi,
University Medical Center
Hamburg-Eppendorf, Germany

***Correspondence:**
Sou Nobukawa
nobukawa@cs.it-chiba.ac.jp

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: 11 October 2021
Accepted: 03 January 2022
Published: 04 February 2022

Citation:
Ando M, Nobukawa S, Kikuchi M and Takahashi T (2022) Alteration of Neural Network Activity With Aging Focusing on Temporal Complexity and Functional Connectivity Within Electroencephalography. Front. Aging Neurosci. 14:793298. doi: 10.3389/fnagi.2022.793298

With the aging process, brain functions, such as attention, memory, and cognitive functions, degrade over time. In a super-aging society, the alteration of neural activity owing to aging is considered crucial for interventions for the prevention of brain dysfunction. The complexity of temporal neural fluctuations with temporal scale dependency plays an important role in optimal brain information processing, such as perception and thinking. Complexity analysis is a useful approach for detecting cortical alteration in healthy individuals, as well as in pathological conditions, such as senile psychiatric disorders, resulting in changes in neural activity interactions among a wide range of brain regions. Multi-fractal (MF) and multi-scale entropy (MSE) analyses are known methods for capturing the complexity of temporal scale dependency of neural activity in the brain. MF and MSE analyses exhibit high accuracy in detecting changes in neural activity and are superior with regard to complexity detection when compared with other methods. In addition to complex temporal fluctuations, functional connectivity reflects the integration of information of brain processes in each region, described as mutual interactions of neural activity among brain regions. Thus, we hypothesized that the complementary relationship between functional connectivity and complexity could improve the ability to detect the alteration of spatiotemporal patterns observed on electroencephalography (EEG) with respect to aging. To prove this hypothesis, this study investigated the relationship between the complexity of neural activity and functional connectivity in aging based on EEG findings. Concretely, MF and MSE analyses were performed to evaluate the temporal complexity profiles, and phase lag index analyses assessing the unique profile of functional connectivity were performed based on the EEGs conducted for young and older participants. Subsequently, these profiles were

combined through machine learning. We found that the complementary relationship between complexity and functional connectivity improves the classification accuracy among aging participants. Thus, the outcome of this study could be beneficial in formulating interventions for the prevention of age-related brain dysfunction.

Keywords: EEG signal, multi-scale entropy, aging, functional connectivity, multi-fracial

1. INTRODUCTION

Complex temporal variability within brain activity plays an important role in perceptual and overall mind and behavioral processes and is known to be a mechanism for stochastic resonance and facilitation (as reviewed in McDonnell and Ward, 2011; Garrett et al., 2013; Takahashi, 2013; Yang and Tsai, 2013; Nobukawa and Nishimura, 2020). Moreover, various memory function components, cognitive functions, and perceptual functions of the brain are associated with brain activity at each temporal scale, as well as with frequency-band specific behaviors, such as theta, beta, alpha, and gamma bands (Klimesch et al., 2007). Therefore, studies using high time resolution for electroencephalography (EEG) and magnetoencephalography (MEG) are currently being conducted to evaluate the complexity of high-frequency components. In particular, neural fluctuations with temporal scale dependency, which can be observed with EEG and MEG, including their relationship with cognitive function (McIntosh et al., 2008), development (Hasegawa et al., 2018), aging (Takahashi et al., 2009, 2016; Nobukawa et al., 2019a), and the pathology of mental disorders (Takahashi et al., 2010; Ahmadlou et al., 2011; Nobukawa et al., 2019b, 2020a), have been extensively studied. Multi-scale entropy (MSE) and multi-fractal (MF) analyses are widely utilized as an effective evaluation method for complexity with temporal scale dependency (as reviewed in Takahashi, 2013; Yang and Tsai, 2013). These methods that focus on the temporal scale dependency of complexity can capture alterations in brain activity within a variety of psychiatric disorders (Yang and Tsai, 2013). In particular, EEG signals in schizophrenia have been reported to be less complex in the frontal region based on a slow temporal scale (Paulus et al., 1996; Takahashi et al., 2010). Similarly, patients with bipolar disorder show less neural complexity (Gottschalk et al., 1995). Considering the temporal scale dependence within EEG for Alzheimer's disease (AD), MSE analysis showed low complexity in the frontal region (Mizuno et al., 2010; Ni et al., 2016). In addition, MF analysis can detect the severity of cognitive impairment in AD (Zorick et al., 2020). Moreover, recent studies have shown that the profile of temporal complexity for EEG signals can be utilized for classifying EEG for AD, and the combination of complexity profiles obtained via MF and MSE enhances the accuracy of AD identification based on their complementary relationship (Zorick et al., 2020; Ando et al., 2021). Consequently, approaches for combining complexity profiles could open new avenues for the identification and characterization of the complex patterns of neural activity regarding cognitive alteration in psychiatric disorders.

In addition to the complex temporal variability, functional connectivity reflects the integration of brain information processes in each neural region, which are represented as mutual interactions of neural activity among brain regions (reviewed in Varela et al., 2001; Buzsáki and Draguhn, 2004; Fries, 2005; Hutchison et al., 2013). Therefore, functional connectivity correlates with cognitive function and alters several pathological conditions characterized by impairments in cognitive function, such as AD (Hata et al., 2016; Yu et al., 2016), autism spectrum disorder (ASD) (Righi et al., 2014), and attention deficit hyperactivity disorder (ADHD) (Ueda et al., 2020). Functional connectivity reflected in EEG has been quantified by coherence, correlation, and mutual information analyses, which reflect the degree of synchronization of neural activity between brain regions (Aertsen et al., 1989; Friston et al., 1993; Bullmore and Sporns, 2009). In recent years, measured values, such as synchronization likelihood (Stam and Van Dijk, 2002) and the phase lag index (PLI) (Stam et al., 2007), have been used as an evaluation method for phase synchronization to solve the problem of volume conduction as a cause for the detection of spurious synchronizations (Nunez et al., 1997; Nolte et al., 2004). By utilizing this advantage of the PLI within EEG, alterations in functional connectivity under pathological conditions have been revealed in previous studies (Engels et al., 2015; Ueda et al., 2020; Nobukawa et al., 2020a). For example, children with ADHD were reported to have a lower gamma PLI than children with typical development (Ueda et al., 2020); AD is associated with a reduced alpha, beta, and gamma PLI compared with that observed in healthy controls (Nobukawa et al., 2020a). Likewise, patients with schizophrenia reportedly demonstrate a reduced PLI of the beta band in the frontal region and a reduced PLI of the gamma band throughout the scalp (Takahashi et al., 2018). The PLI has also been used to assess frequency dependence in children with ASD (Takahashi et al., 2017). Furthermore, the PLI can capture functional connectivity within high cognitive functions among healthy older participants (Nobukawa et al., 2020b). PLI is robust against artifacts such as body and eye movements, and muscle activation thus the influence of artifacts on PLI is relatively small, because the major parts of this influence lie in the amplitude space of signals, while PLI estimates phase-based functional connectivity (Stam et al., 2007). However, in the higher frequency gamma band range, artifacts due to muscle activity are larger compared to slower frequency ranges (Whitham et al., 2007, 2008); therefore, there may be issues with PLI estimation accuracy in the gamma range (Lau et al., 2012; Engels et al., 2015).

In recent trends within neural activity analysis, multiple spatio-temporal profiles of neural activity (which combine profiles obtained by several evaluation methods) are integrated

via machine learning; subsequently, analyses detecting the pathology of several psychiatric disorders and estimating the ability of brain function have been conducted with higher accuracy as compared with using a single profile (reviewed in Vu et al., 2018). In particular, informative studies have been conducted combining profiles of functional connectivity and temporal complexity (Ghanbari et al., 2015; Nobukawa et al., 2020a). Studies have also reported a complementary relationship between functional connectivity and neural complexity (Ghanbari et al., 2015; Nobukawa et al., 2020a). In patients with ASD, increasing (or decreasing) complexity decreases (or increases) functional connectivity, suggesting that the functional connectivity and complexity are complementary (Ghanbari et al., 2015). For patients with AD, the relationship between functional connectivity and complexity shows different temporal scales and region-specific dependencies in both healthy participants and among patients with AD, suggesting that the relationship between functional connectivity and complexity may reflect the complex pathological process occurring within AD (Nobukawa et al., 2020a). However, to the best of our knowledge, an approach combining functional connectivity and the complexity of neural activity has not been evaluated under healthy conditions. Even in the healthy aging process, brain functions, such as attention, memory, and cognitive functions, degrade over time (Birren and Fisher, 1995). Therefore, in a super-aging society, the alteration of spatial-temporal neural activity owing to aging is considered crucial for interventions for the prevention of brain dysfunction.

Thus, we hypothesized that the complementary relationship between functional connectivity and complexity could improve the ability to detect alteration of spatiotemporal patterns within EEG with respect to the aging process. In this study, MF and MSE analyses were performed to evaluate the temporal complexity profiles, and PLI analyses evaluating the unique profile of functional connectivity were performed based on EEG among younger and older participants. Subsequently, these profiles were combined via machine learning methodology.

2. MATERIALS AND METHODS

2.1. Participants

A total of 32 healthy younger people (15 males, 17 females; average age, 23.9 years; standard deviation [SD], 4.7 years; age range, 20–35 years) and 18 healthy older people (7 males, 11 females; average age, 57.5 years; SD, 4.7 years; age range, 51–67 years) were enrolled in this study. These groups were sex-matched ($\chi^2 = 0.30$, $p = 0.59$). The older participants were all non-smokers and were not on any medications. Participants with medical or neurological conditions (including epilepsy or head trauma occurring in the past), as well as those with a history of alcohol or drug dependence, were excluded from the current study. All the participants provided their written informed consent following an explanation of study procedures as well as risks and benefits by study personnel. This study was approved by the Ethics Committee of Kanazawa University and was conducted in accordance with the principles of the Declaration of Helsinki and its later amendments. The EEG data used in this study evaluated the dynamics of phase synchronization (Nobukawa et al., 2019a).

2.2. EEG Recordings

Methods for recording and pre-processing EEG data have been reported and established in previous research (Mizuno et al., 2010). Specifically, the participants in the current study sat in a soundproof recording room, and their EEG was measured under controlled room lighting conditions. For EEG measurement, 16 electrodes (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, Fz, Pz, T5, and T6) were used; this system was based on the recommended electrode arrangement under the international 10–20 system. EEG activity was measured with reference to the binaural connection. The EEG-4518 monitor used for electroencephalogram measurements in this study was manufactured by Nihon Kohden Co., Ltd. (Tokyo, Japan). The sampling frequency was 200 Hz for the recording. The electrode/skin conductance impedance was controlled to within less than 5kΩ for each electrode. Participants' electroencephalogram signals were measured for 10–15 min in a resting state with the eyes closed. Researchers visually inspected the participants' arousal using a video surveillance system; participants were asked to close their eyes, and researchers confirmed that only awake epochs were measured. If the alpha and theta oscillations became weaker or stronger, compared with ones at beginning stage of the recording, this duration was not used for evaluation, because this duration belonged to the light sleep stage. Additionally, the EEG signals were visually assessed to identify artifacts, such as muscle activity, blinks, and eye movement; consequently, 60-s (12,000 data points) of artifact-free time-series segments within the EEG signals recorded in the awake state with eyes closed were identified. For each epoch, bandpass filtering with the range of 2.0–60 Hz was applied. The first and last 5-s period (1,000 data point) in each bandpass-filtered epoch were removed to avoid transient behaviors produced by the bandpass filtering process. MSE and MF analyses were performed for 50 consecutive seconds (i.e., 10,000 data points) of epochs. In the PLI analyses, values decrease with increasing epoch length (Fraschini et al., 2016); therefore, it is difficult to identify changes with an increasing epoch length. In addition, using short epoch lengths makes it impossible to capture information on slow frequency components. To balance these considerations, the PLI analysis divided 50 consecutive seconds (10,000 data points) into 10 epochs of 5 s each (Takahashi et al., 2017, 2018; Nobukawa et al., 2020a,b).

2.3. Multi-Fractal Analysis

The overview of flow for multi-fractal analysis is shown in **Figure 1A**. In MF analyses, wavelet leaders derived from the coefficients of the discrete wavelet transform are widely used (Jaffard et al., 2006; Wendt and Abry, 2007). MF analysis is an analysis method that uses the Hölder index to represent the fractal dimension of the partial structure that characterizes

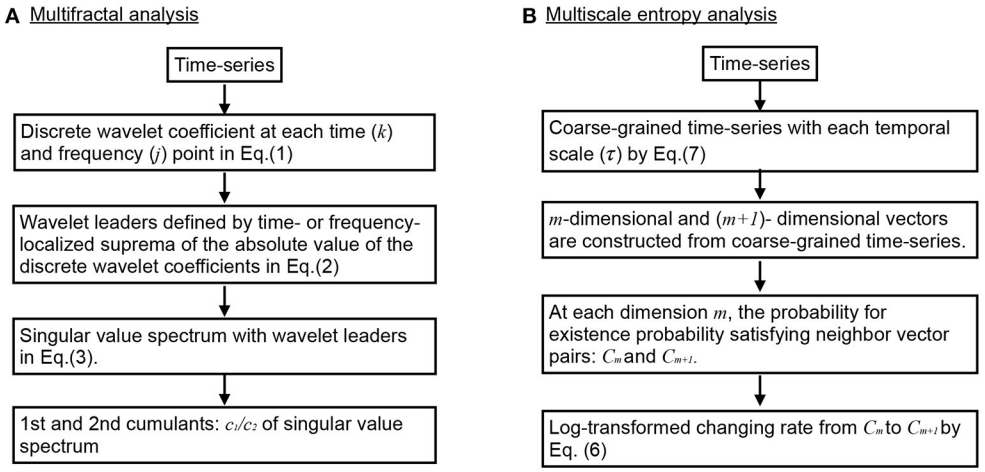


FIGURE 1 | Overview of flow for complexity analysis. **(A)** Multifractal (MF) analysis. **(B)** Multiscale entropy (MSE) analysis.

the structure of data X via spectrum data. The discrete wavelet coefficient of the discrete signal $X(t)$ is given by Equation (1).

$$d_X(j,k) = \int_R X(t)2^j, \psi_0(2^{-j}t - k)dt \quad (j = 1, 2, ..., k = 1, 2, ...) \tag{1}$$

Here, ψ_0 is a compact-supported mother wavelet function. The Equation (2) shows one-dimensional wavelet leaders which are time- or frequency-localized suprema of the absolute value of the discrete wavelet coefficients $d_X(j,k)$:

$$L_x(j,k) = \sup_{\lambda' \subset 3\lambda_{j,k}} |d_X(\lambda')| \tag{2}$$

Here, $\lambda = \lambda_{j,k} = [k2^{-j}, (k+1)2^{-j}]$ represents the time interval of the scale 2^{-j}. Additionally, $3\lambda_{j,k-1} = \cup\lambda_{j,k} \cup \lambda_{j,k+1}$ represents the adjacent time (Wendt and Abry, 2007). The spectrum of singularity of L_X is defined by Equation (3) with wavelet leaders (Jaffard et al., 2006; Wendt and Abry, 2007).

$$D(h) = \inf_{q \neq 0}(1 + qh - \zeta_L(q)) \tag{3}$$

Here, h indicates the Hölder index. Also, q indicates the moment of the wavelet leaders. The scaling index $\zeta_L(q)$ is defined by Equation (4). The wavelet leader structure function $S_L(q,j)$ is defined by Equation (5).

$$\zeta_L(q) = \liminf_{j \to \infty} \left(\frac{\log_2 S_L(q,j)}{\log_2 2^{-j}} \right) \tag{4}$$

$$S_L(q,j) = \frac{1}{n_j} \sum_{k=1}^{n_j} |L_X(j,k)|^q \tag{5}$$

Here, n_j indicates the number of samples of X when the scale is 2^j. As the Hölder index h approaches 1.0, the time-series shape becomes more differentiable. However, as the Hölder index h approaches 0, the time-series shape becomes nearly discontinuous. A signal is monofractal if the scaling index $\zeta_L(q)$ is a linear function and $D(h)$ converges to a particular h. Contrastingly, the fact that the signal is multi-fractal indicates a scaling index, where $\zeta_L(q)$ deviates from linearity and $D(h)$ is widely distributed in h. In this study, to capture the profile of $D(h)$, the primary cumulant c_1 of $D(h)$ was used as an indicator of the smoothness of the entire time series signal, and the secondary cumulant c_2 was used as an index evaluating the local fluctuation of the time-series signal. For the multi-fractal time series, $D(h)$ is distributed around c_1. Therefore, the degree of distribution of $D(h)$ reflects the multi-fractal property, which corresponds to $|c_2|$. The time-series with large (small) multi-fractality ($|c_2|$) exhibits intermittent and transient behavior with large (small) amplitude (Ihlen, 2012); while, the complexity notified by c_1 reflects the degree of complexity for temporal behavior in entire time-range, instead of intermittent behavior (see **Supplementary Material**). In this study, multi-fractal analysis was performed using the wavelet toolbox in MATLAB (https://jp.mathworks.com/products/wavelet.html; MathWorks, Natick, MA, USA).

2.4. Multi-Scale Entropy Analysis

The overview of flow for MSE analysis was shown in **Figure 1B**. MSE analysis was used to assess the temporal scale dependence for EEG time series complexity (Costa et al., 2002). The time-series sample entropy of the random Z-score variables $\{x_1, x_2, ..., x_N\}$ is given by Equation (6).

$$h(r,m) = -\log \frac{C_{m+1}(r)}{C_m(r)}. \tag{6}$$

Here, $C_m(r)$ is the probability of $|\mathbf{x}_i^m - \mathbf{x}_j^m| < r (i \neq j, i, j = 1, 2, ...)$ among all pairs of i and j. $\mathbf{x}_i^m$ indicates an m-dimensional vector

$\mathbf{x}_i^m = \{x_i, x_{i+1}, ..., x_{i+m-1}\}$. In the MSE analysis, $\{x_1, x_2, ..., x_N\}$ is calculated using Equation (7) for coarse-grained time series y_j.

$$x_j = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j_\tau} y_i (1 \leq j \leq \frac{N}{\tau}). \tag{7}$$

Here, $\{y_1, y_2, ..., y_N\}$ represents the observed signals. $\tau(\tau = 1, 2, ...)$ represents the temporal scale. In this study, we set $m = 2$ and $r = 0.2$ (Costa et al., 2002) and MSE analysis was performed using the Physio Toolkit toolbox in MATLAB (http://physionet.incor.usp.br/physiotools/sampen/).

2.5. Phase Lag Index Analysis

The PLI was obtained to measure phase synchronization, and the characteristics of the synchronization signal were quantitatively estimated. The EEG signal was divided into five frequency bands: the delta (2–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma bands (30–60 Hz). Here, several studies showed that artifacts in the gamma band due to muscle activity is larger compared with slower frequency bands (Whitham et al., 2007, 2008). Therefore, through visual examination of EEG signals, the time including muscle activity was avoided in the evaluation epochs. Each band division divides the signal at time t, and the point ϕ_a is indicated by phase $\phi_a(t)$ and amplitude $Aa(t)$ using the Hilbert transform. In addition, the phase difference $\Delta\phi_{ab}(t_i)$ observed between signals with two different points a and b at time t_i is given by Equations (8) and (9) (Stam et al., 2007).

$$\Delta\phi_{ab}(t_i) = \phi_a(t_i) - \phi_b(t_i) \tag{8}$$

$$\Delta\phi_{\mathrm{mod}}(t_i) = \Delta\phi_{ab}(t_i) \mathrm{mod} 2\pi \tag{9}$$

The PLI of the signal between the two points a and b for the duration T is given by Equation (10).

$$PLI_{ab} = \left| \frac{1}{T} \sum_{i=0}^{T} \mathrm{sign}(\Delta\phi_{\mathrm{mod}}(t_i)) \right| \tag{10}$$

When signals with the same source are observed at different points, $\Delta\phi_{ab}(t_i)$ is 0 and $\Delta\phi_{\mathrm{mod}}(t_i) = 0$; subsequently the PLI_{ab} value becomes 0. In addition, the observation at the point on the opposite side of the electric dipole is defined as $\Delta\phi_{ab}(t_i) = \pi$ within Equation (8) in cases where the signal source is assumed to follow the dipole model. This yields $PLI_{ab} = 0$. The average PLI of any electrode a via another electrode $b = 1, 2, ..., K(b \neq a)$ (called the node strength; NS) is given by Equation (11). Here, K represents the total number of electrodes $K = 16$.

$$NS_a = \frac{1}{K-1} \sum_{b=1, b\neq a}^{K} PLI_{ab} \tag{11}$$

2.6. Statistical Analysis

For c_1 and $|c_2|$, repeated-measures analysis of variance (ANOVA) was performed to determine statistically significant differences between the younger and older groups. Age group was used as an inter-subject factor, and the 16 electrodes from Fp1 to T6 were used as intra-subject factors. The ANOVA results were represented by F values based on intra-group and inter-group variance comparisons. Greenhouse-Geisser adjustments were applied to the degrees of freedom. The $\alpha = 0.05$ bilateral level was used; this was considered a statistically significant criterion for avoiding type I errors. A *post-hoc t*-test was subsequently used to evaluate the main effect between the younger and older age groups and effect of the interactions per electrode. Here, Benjamini-Hochberg false discovery rate (FDR) correction was applied to the t value for multiple comparisons of c_1 and $|c_2|$ ($q < 0.05$) (16 p values: 16 electrodes).

In the ANOVA for sample entropy, age group was used as an inter-subject factor, and the 16 electrodes from Fp1 to T6 and a temporal scale were used as intra-subject factors. A *post-hoc t*-test was subsequently used to evaluate the main effect between the younger and older groups and effects of interaction for the electrodes and temporal scales. The $\alpha = 0.05$ bilateral level was used. FDR correction was applied to the t scores for multiple comparisons ($q < 0.05$) (480 p values: 16 electrode × 30 scales).

In ANOVA for NS at each frequency band, age group was used as an inter-subject factor, and the 16 electrodes (from Fp1 to T6) were used as intra-subject factors. A *post-hoc t*-test was subsequently used to evaluate the main effect between the younger and older groups and effect of the interaction for the electrodes. The $\alpha = 0.05$ bilateral level was used. FDR correction was applied to the t scores for multiple comparisons ($q < 0.05$) (80 p values: 16 electrodes × 5 frequency bands). The t-test was used for electrode-pair-wise group comparison of PLI between the younger and older groups. With a control for multiple comparisons, FDR correction was applied to the t scores ($q < 0.05$) (600 p values: 120 electrode pairs × 5 bands).

Older participants were classified using the receiver operating characteristics (ROC) curve. A logistic regression model based on sample entropy, c_1, $|c_2|$, and the NS of the PLI was used to identify older participants. Here, the logistic regression model outputs the "older participants" discrimination probability for each participant. The true/false positive rate at each threshold of discrimination probability from 0 to 1.0 for both groups was then measured. Principal component analysis was used as a pre-treatment for dimensionality reduction, and logistic regression based on c_1, $|c_2|$, sample entropy, and the NS of the PLI was implemented. The accuracy of discrimination was evaluated using the area under the ROC curve (AUC). We also used 5-fold cross-validation to prevent overfitting; AUC= 1.0 corresponds to perfect discrimination, and AUC= 0.5 corresponds to random discrimination. Here, the principal component analysis was conducted within cross validation (Shim et al., 2021) to avoid the inaccurate estimation of performance of discrimination. AUC values were averaged among 20 trials to choose tested and evaluated data set in 5-fold cross-validation and their standard deviations (SD) were also derived.

To evaluate the relationship between NS and complexity, we used Spearman's correlation R between NS and the complexity indexes ($c_1, |c_2|$). To control the multiple comparison, FDR correction was applied these R-scores ($q < 0.05$) (16 p values: 16 electrodes).

3. RESULTS

3.1. Multi-Fractal Analysis

MF analysis was performed in the younger and older participants. **Figure 2** shows the mean and standard deviation for each group with respect to $D(h)$ and h. Owing to its wide distribution, this analysis is thought to reflect the multi-fractal characteristics (Sikdar et al., 2018) of the EEG signals for both groups. **Table 1** shows the repeated-measures ANOVA results of the first (c_1) and second ($|c_2|$) cumulants within a singular spectrum. A strong main effects were observed for c_1 and $|c_2|$. The mean values of c_1 and $|c_2|$ in the older and younger groups, as well as the results of a *post-hoc* t-test between the older and younger groups, are shown in **Figure 3**. The *post-hoc* t-test revealed that the value of c_1 was statistically significantly lower for the older participants at 13 electrodes (F3, Fz, F4, F7, F8, C3, C4, P3, Pz, P4, T6, O1, and O2) ($q < 0.050$). In addition, the results showed that the value of $|c_2|$ was statistically significantly lower for the older participants at 14 electrodes (Fp1, Fp2, F3, Fz, F4, F7, F8, C3, C4, Pz, P4, T6, O1, and O2) ($q < 0.050$). The results of the MF analysis demonstrated that aging increases complexity (shown by less smoothness) and decreases multi-fractality. The time-series with large (small) multi-fractality exhibits intermittent and transient behavior with large (small) amplitude (Ihlen, 2012). Meanwhile, complexity reflects the degree of complexity for temporal behavior in entire the time-range, rather than intermittent behavior. Therefore, EEG signal in older subjects corresponds to homogeneous and highly complex temporal behaviors.

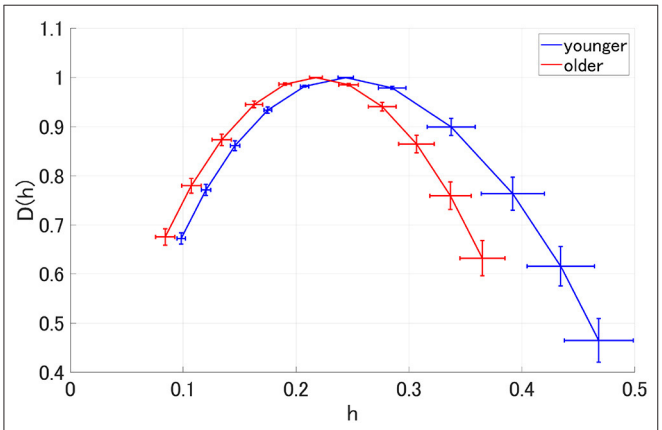


FIGURE 2 | Spectrum of singularity for the older and younger groups. This figure shows the mean and standard deviation of each group for $D(h)$ and h, respectively. Owing to the wide distribution observed here, these results are thought to reflect the multifractal characteristics of the EEG signals for both groups.

3.2. Multi-Scale Entropy Analysis

MSE analysis was performed in the younger and older participants. **Table 2** shows the repeated-measures ANOVA results for the MSE analysis. We found that no main effect was observed, although there were interactions in the group × scale and the group × node × scale. The mean values of sample entropy in the older and younger groups, as well as the results of a *post-hoc* t-test between the older and younger groups, are shown in **Figure 4**. The results demonstrated a statistically significantly higher sample entropy for the older participants ($q < 0.050$) in the temporal-scale region of 1 to 5 (0.005 to 0.025 s at all electrodes). The results of the MSE analysis demonstrated that aging increases complexity on a fast temporal scale.

3.3. Phase Lag Index Analysis

PLI analysis was performed on younger and older participants. **Table 3** shows the ANOVA analysis results for the NS of the PLI for each band among younger and older participants. The results indicated that there was a main effect in the delta and gamma bands and that there was an interaction with respect to the group × node in the alpha, beta, and gamma bands. The *post-hoc* t-test results for the NS are shown in **Figure 5C**. Although no statistically significant differences satisfying with FDR criteria $q < 0.05$ were observed between the older and younger groups, relatively higher NS at delta and gamma band in the older group was observed. Regarding the PLI among pair-wise electrodes, the mean values of the PLI in the older and younger groups, as well as the results of t-tests between the older and younger groups, are shown in **Figures 5A,B**. No statistically significant differences satisfying with FDR criteria ($q < 0.05$) were observed between the older and younger groups.

3.4. Correlation Analysis Between Complexity and Functional Connectivity

To evaluate the relationship between complexity and functional connectivity, a correlation analysis was performed, using

TABLE 1 | Younger vs. older repeated measure ANOVA analysis results [F value (p value, partial η^2)] in multifractal (MF) analysis.

	Group	Group × nodes	Degree of freedom (ϵ)
c_1	**$F = 25.25$ ($p < 0.001$, $\eta^2 = 0.345$)**	$F = 1.73$ ($p = 0.13$, $\eta^2 = 0.035$)	5.06 ($\epsilon = 0.034$)
$\|c_2\|$	**$F = 22.23$ ($p < 0.001$, $\eta^2 = 0.317$)**	$F = 1.73$ ($p = 0.11$, $\eta^2 = 0.035$)	6.04 ($\epsilon = 0.035$)

F and p values with $p < 0.05$ are represented by bold characters. Degree of freedom and Greenhouse-Geisser adjustments ϵ in the interaction for group × nodes are also shown.

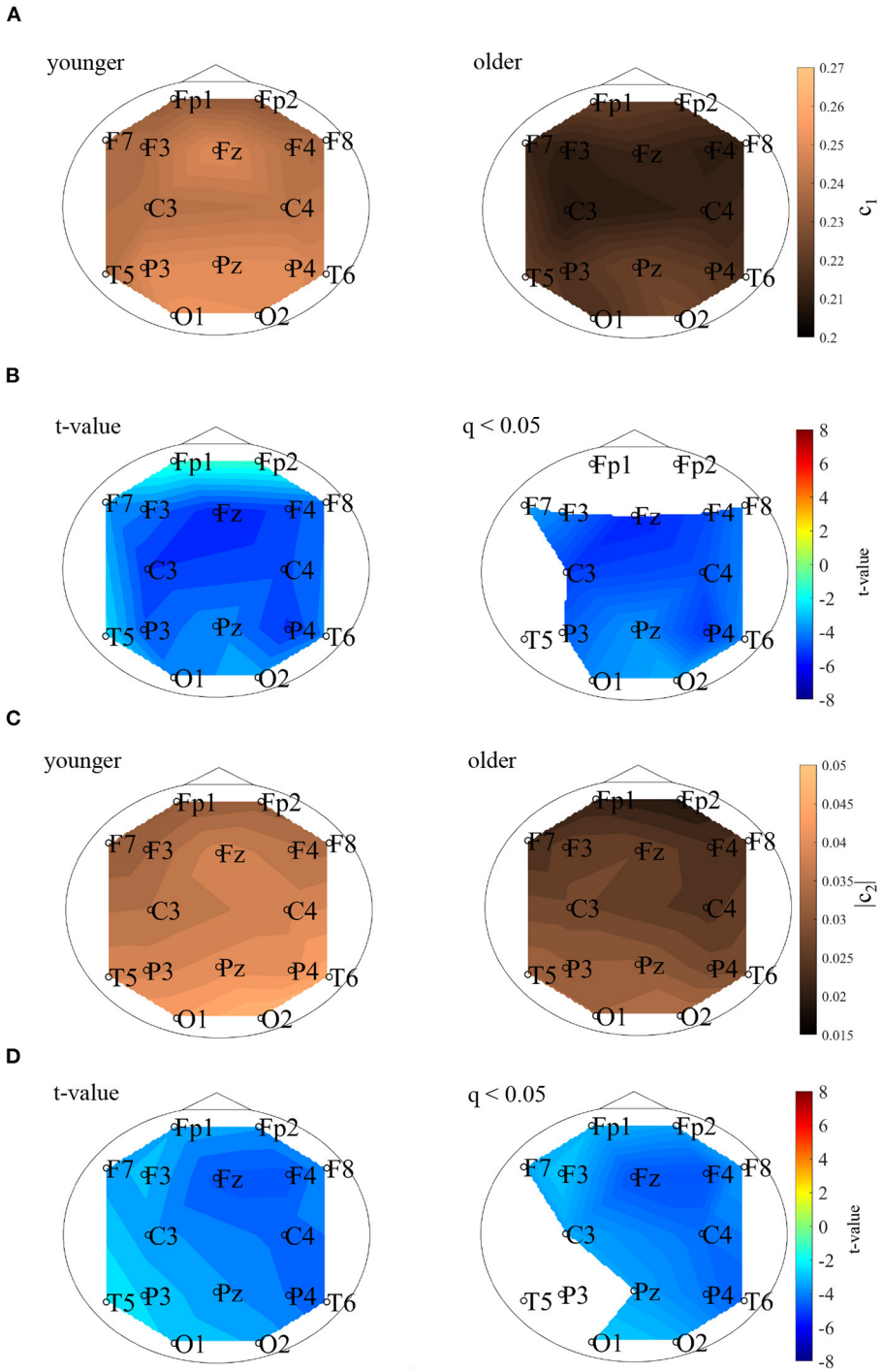


FIGURE 3 | (A) 1st cumulant of the spectrum of singularity c_1. The mean values of c_1 in the younger (left) and older (right) groups. **(B)** *t* values comparing the younger and older groups. Warm (cold) colors represent higher (smaller) c_1 values for older versus younger participants. The left and right of the figure correspond to the *t*- and *t*-values satisfying the false discovery rate (FDR) correction criteria $q < 0.050$. The c_1 value for the older group, which had statistically significantly lower values at F3, Fz, F4, F7, F8, C3, C4, P3, Pz, P4, T6, O1, and O2, and is shown here. **(C)** Absolute value of 2nd cumulant of the spectrum of singularity $|c_2|$. The mean values of $|c_2|$ in the younger (left) and older (right) groups are shown here. **(D)** *t*-values comparing the older and younger groups. Warm (cold) colors represent higher (smaller) $|c_2|$ values for older versus younger participants. The left and right correspond to the *t*- and *t*-values satisfying the FDR correction criteria $q < 0.050$. The $|c_2|$ of the older group had statistically significantly lower values at Fp1, Fp2, F3, Fz, F4, F7, F8, C3, C4, Pz, P4, T6, O1, and O2.

TABLE 2 | Younger vs. older repeated–measures ANOVA results [F value (p value, partial η^2)], degree of freedom and Greenhouse-Geisser adjustments ϵ in multi scale entropy (MSE) analysis.

Group	Group × node	Group × scale	Group × node × scale
$F = 3.37$ ($p = 0.073$, $\eta^2 = 0.066$)	$F = 1.51$ ($p = 0.21$, $\eta^2 = 0.030$, degree of freedom: 3.62, $\epsilon = 0.242$)	**$F = 19.93$ ($p < 0.001$, $\eta^2 = 0.293$ degree of freedom: 2.94, $\epsilon = 0.102$)**	**$F = 1.949$ ($p = 0.020$, $\eta^2 = 0.039$ degree of freedom: 13.898, $\epsilon = 0.032$)**

F and p values with $p < 0.05$ are represented by bold characters.

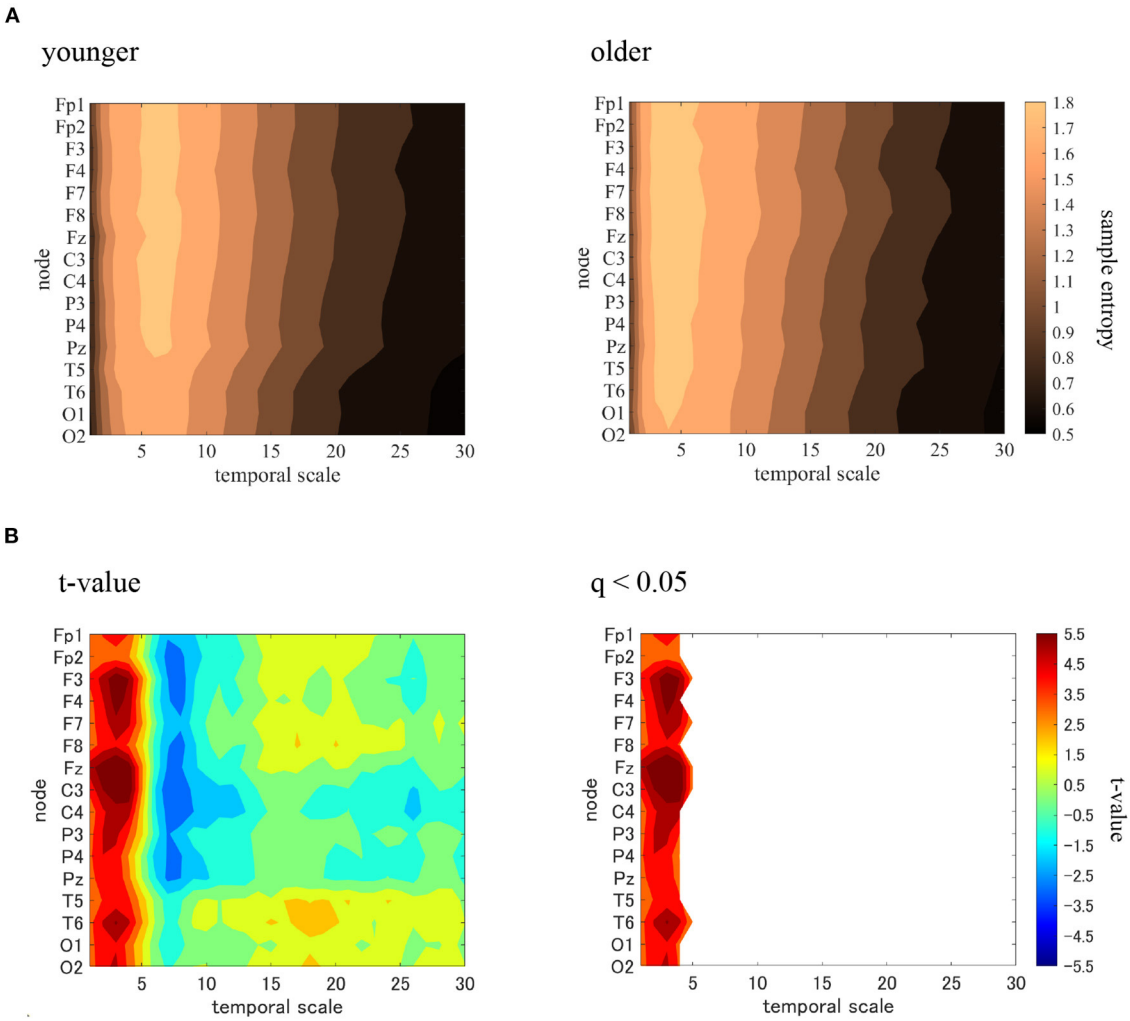


FIGURE 4 | Multi-scale entropy analysis in younger and older groups. The horizontal axis represents the temporal-scale factor, τ. **(A)** Mean values of sample entropy from 1 (0.005 s) to 30 (0.15 s) scale factors in younger (left part) and older (right part) participants are shown here. **(B)** t-values comparing the older and younger groups are shown here as well (left part). Warm (cold) colors represents a higher (smaller) sample entropy value for older individuals than for younger participants. The observed t-value satisfies the FDR correction criterion $q < 0.050$ (right part). A statistically significantly higher sample entropy of the low temporal-scale regions 1 to 5 (0.005–0.025 s) is depicted here.

Spearman's correlation, between NS and the complexity indexes ($c_1, |c_2|$). **Figure 6** shows the Spearman's correlations between NS of PLI and c_1 and between NS of PLI and $|c_2|$ in both the younger and older groups. The correlation with c_1 did not meet the FDR correction criteria of $q < 0.050$; while there were positive correlations with NS at alpha and $|c_2|$ at Fp1, Fp2, F3, F4, and Fz as well as NS at beta band and $|c_2|$ at F3 in younger group. In **Figure 7**, the scatter plots at these electrodes were shown, significantly large positive correlations were observed. Therefore, large node strength might lead the intermittent and transient behavior reflecting $|c_2|$ in local neural activity, instead of steady neural variability reflecting c_1.

TABLE 3 | Younger vs. older repeated-measure ANOVA analysis results [F value (p value, partial η^2)] in the node strength (NS) of phase lag index (PLI).

	Group	Group × nodes	Degree of freedom (ϵ)
delta	$\mathbf{F = 4.18\ (p = 0.046, \eta^2 = 0.80)}$	$F = 1.26\ (p = 0.262, \eta^2 = 0.026)$	$8.183\ (\epsilon = 0.546)$
theta	$F = 0.04\ (p = 0.833, \eta^2 = 0.001)$	$F = 0.98\ (p = 0.453, \eta^2 = 0.020)$	$8.948\ (\epsilon = 0.597)$
alpha	$F = 1.09\ (p = 0.301, \eta^2 = 0.022)$	$\mathbf{F = 3.95\ (p < 0.001, \eta^2 = 0.076)}$	$5.899\ (\epsilon = 0.393)$
beta	$F = 0.097\ (p = 0.757, \eta^2 = 0.002)$	$\mathbf{F = 2.89\ (p = 0.006, \eta^2 = 0.057)}$	$7.060\ (\epsilon = 0.471)$
gamma	$\mathbf{F = 4.17\ (p = 0.047, \eta^2 = 0.080)}$	$\mathbf{F = 3.35\ (p = 0.002, \eta^2 = 0.065)}$	$6.785\ (\epsilon = 0.452)$

F and p values with $p < 0.05$ are represented by bold characters. Degree of freedom and Greenhouse-Geisser adjustments ϵ in the interaction for group × nodes are also shown.

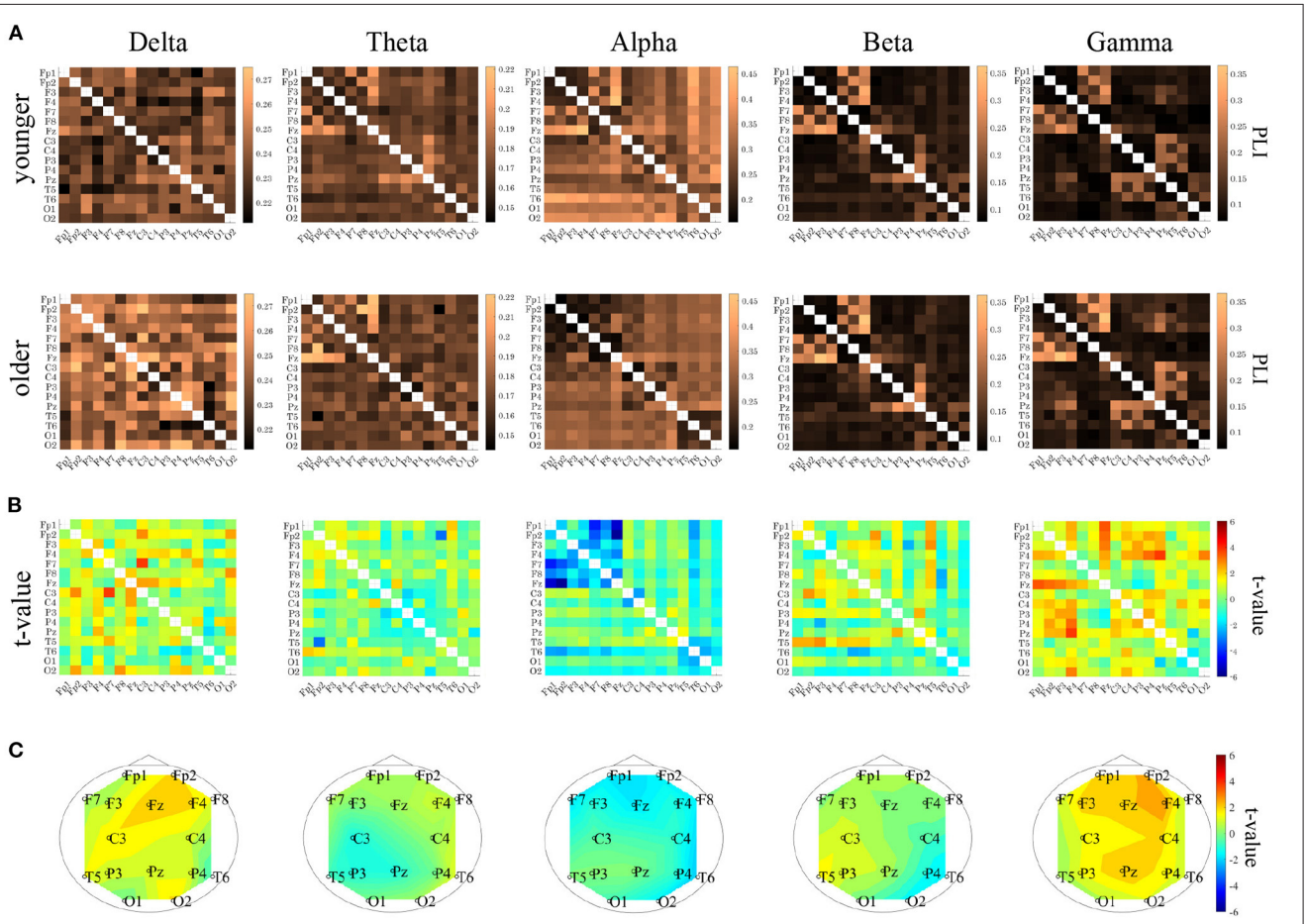


FIGURE 5 | (A) Mean values of phase lag index (PLI) for each band in younger (upper parts) and older (lower parts) groups. **(B)** t-values between the younger and older groups (upper parts). The warm (cold) colors represents higher (smaller) PLI values in the older group than in the younger group. (lower parts). **(C)** t-values for the node strength (NS) of PLI between the older and younger groups. Warm (cold) colors represent a higher (smaller) NS for older versus younger individuals. Although no statistically significant differences satisfying with FDR criteria $q < 0.05$ were observed between the older and younger groups, relatively higher NS at delta and gamma band in the older group was observed. In PLI among pair-wise electrodes, no statistically significant differences satisfying with FDR criteria ($q < 0.05$) were observed between the older and younger groups.

3.5. ROC Curve Analysis

To evaluate the classification ability for c_1, $|c_2|$, and the PLI, we evaluated the ROC. We observed a statistically significantly large sample entropy in the older group for time scales of 1 to 5, as shown in **Figure 4**. Therefore, the sample entropy was averaged in this temporal region for the purpose of clarification. **Table 4** shows the results of the ROC in cases with the first-third principal components of c_1 and $|c_2|$, as well as sample entropy. In the results shown in **Table 4**, c_1 had the highest value (AUC = 0.86). **Table 5** shows the results of the ROC in cases with the first-third principal components for the NS of the PLI in the delta, theta, alpha, beta, and gamma bands. In

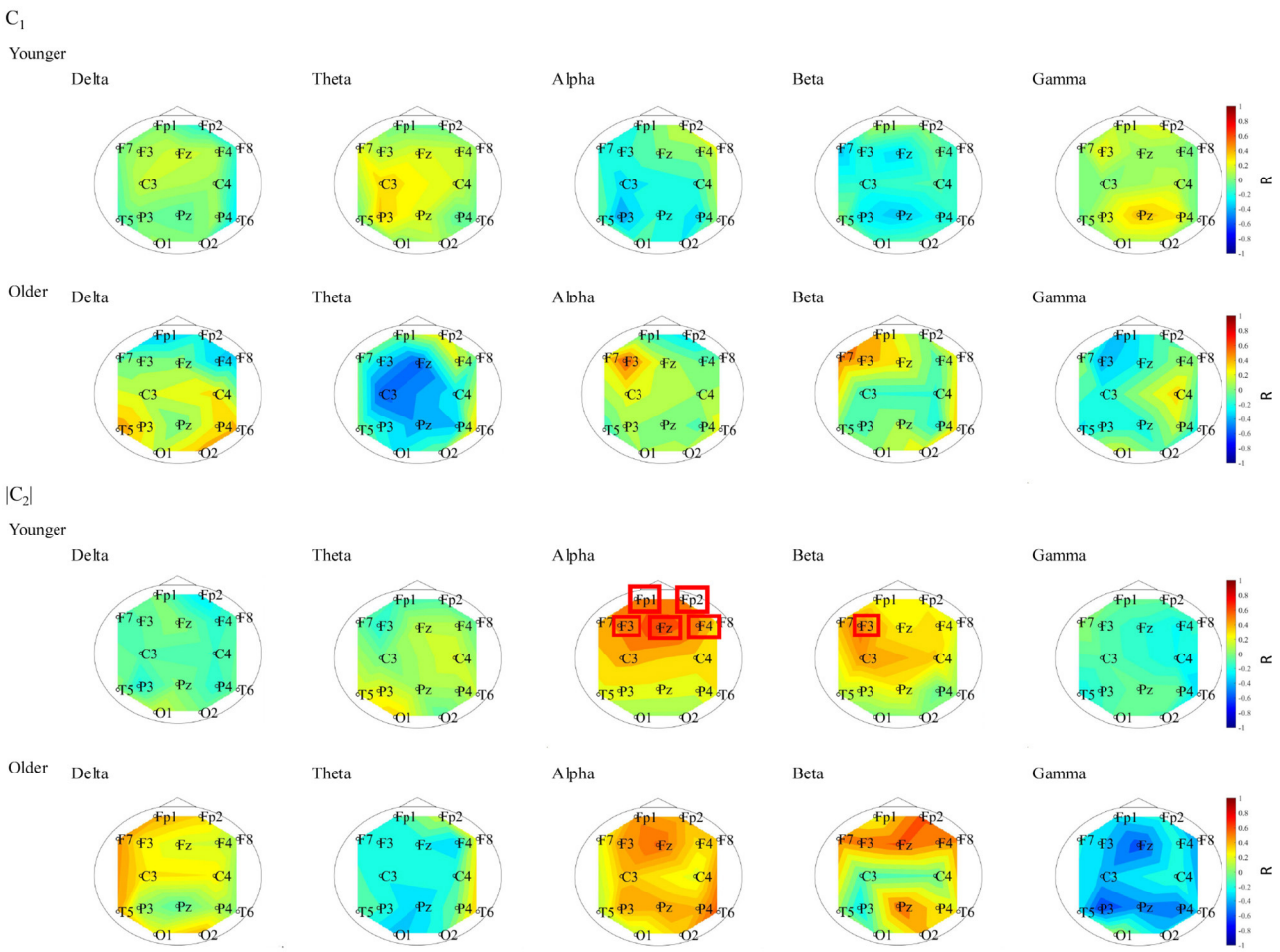


FIGURE 6 | Spearman's rank correlation coefficient R between NS of PLI and c_1 in younger (1st line) and older (2nd line) groups. The correlation values did not meet the FDR correction criteria of $q < 0.050$. Spearman's rank correlation coefficient R between NS of PLI and $|c_2|$ in younger (3rd line) and older (4th line) groups. The electrodes with correlation values to meet the FDR correction criteria of $q < 0.050$ are surround with a line. In younger group, positive correlation with NS at alpha and $|c_2|$ at Fp1, Fp2, F3, F4, and Fz and positive correlation with NS at beta band and $|c_2|$ at F3 were confirmed.

the results shown in **Table 5**, NS at alpha had the highest value (AUC = 0.84). **Table 6** shows the results of the ROC in cases with the first-third principal components of the combination of c_1 and $|c_2|$ and the combination of the NS at alpha, as well as c_1, $|c_2|$ and sample entropy. Almost AUC values increased by combining the NS at alpha, c_1, $|c_2|$, and sample entropy in comparison with cases using a single index, as shown in **Tables 4**, **5**. While, at the other bands, AUCs in the NS are significantly lower in comparison with the complexity indexes (c_1, c_2, and sample entropy) (see **Tables 4**, **5**). Therefore, AUCs in the case with combinations of NS at the other band and the complexity index are inferior to AUCs in the case using a single complexity index.

To demonstrate that the determination area for older participants is determined by c_1, $|c_2|$, sample entropy, and the NS at alpha of the PLI, the determination area of the older participants was defined as the plane between the first principal components of c_1 and first principal components of $|c_2|$, and plane between the first principal components of c_1 and first principal components of the NS at alpha, plane between the first principal components of $|c_2|$ and first principal components of the NS at alpha, and plane between the first principal components of sample entropy and first principal components of the NS at alpha (see **Figure 8**). Here, the other components, with the exception of the axis of the planes, were set to average among participants in both the younger and older study groups. The dependency on all of these factors in the decision region was confirmed.

4. DISCUSSION

In this study, we investigated the relationship between complexity and functional connectivity in aging via EEG. In the MF analysis, we found that c_1 (as the index for the smoothness of the EEG signal) decreased with aging, and $|c_2|$

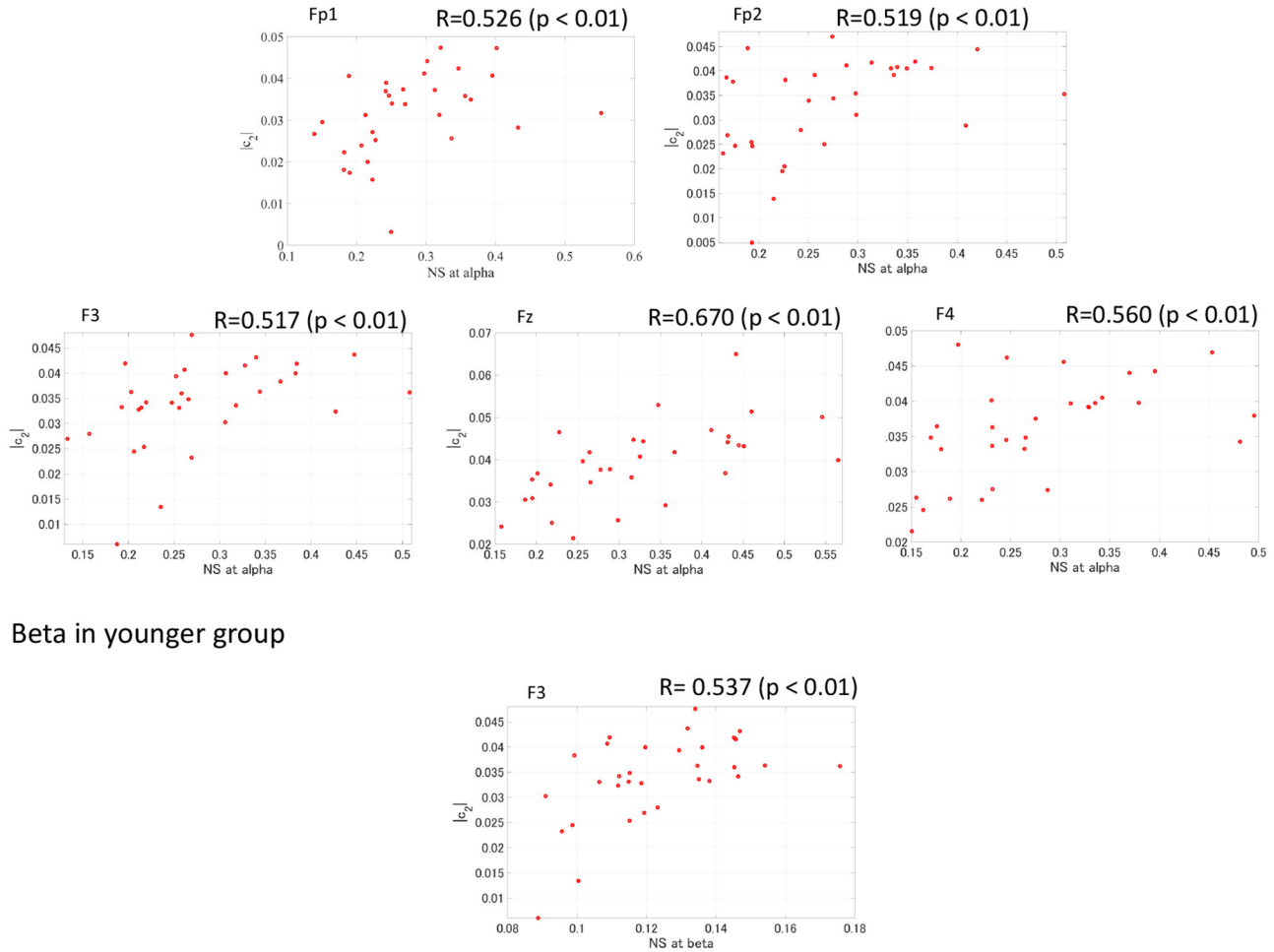


FIGURE 7 | Scatter plot between NS of PLI at alpha/beta bands and $|c_2|$ in younger group where Spearman's rank correlation *R* coefficient satisfies with FDR correction criteria of $q < 0.050$ in **Figure 6**. The significant large positive correlations were observed.

TABLE 4 | The area under the ROC curve (AUC) for c_1, $|c_2|$, and sample entropy averaged scale 1–5.

	AUC (SD)
c_1	0.862 (0.029)
$\|c_2\|$	0.857 (0.026)
sample entropy	0.850 (0.026)

In this case, c_1, $|c_2|$, and sample entropy, each first-third principal components, were used separately. Here, AUC values were averaged among 20 trials to choose tested and evaluated data set in 5-fold cross-validation and their standard deviations (SD) were also derived.

TABLE 5 | AUC for the NS of the PLI. In this case, the NS of the PLI was used separately for each of the first-third principal components.

	AUC (SD)
NS at delta	0.600 (0.055)
NS at theta	0.545 (0.053)
NS at alpha	0.840 (0.030)
NS at beta	0.785 (0.034)
NS at gamma	0.708 (0.058)

Here, AUC values were averaged among 20 trials to choose tested and evaluated data set in 5-fold cross-validation and their standard deviations (SD) were also derived.

(which is an index of a multi-fractal nature) also decreased with aging. In the MSE analysis, a statistically significant region-specific increase in the small-temporal-scale sample entropy of aging was observed. In the PLI analysis, we found that functional connectivity increased in the delta and gamma bands with aging. In the comparison of the classification accuracy among c_1, $|c_2|$, small-temporal-scale sample entropy, and the NS of the PLI, c_1 demonstrated the highest classification accuracy (AUC = 0.86). Considering the complementary

TABLE 6 | AUC for the combination of c_1 and $|c_2|$ and the combination of the NS at alpha, c_1, $|c_2|$ and sample entropy.

	AUC (SD)	*t*-value (*p*-value) with c_1	*t*-value (*p*-value) with $|c_2|$	*t*-value (*p*-value) with sample entropy	*t*-value (*p*-value) with NS at alpha
c_1 & $|c_2|$	0.885 (0.038)	**$t = 2.69$ ($p = 0.014$)**	**$t = 2.96$ ($p = 0.007$)**	-	-
c_1 & NS at alpha	0.881 (0.046)	$t = 1.82$ ($p = 0.083$)	-	-	**$t = 4.82$ ($p < 0.001$)**
$|c_2|$ & NS at alpha	0.887 (0.035)	-	**$t = 3.31$ ($p = 0.003$)**	-	**$t = 5.79$ ($p < 0.001$)**
sample entropy & NS at alpha	0.873 (0.031)	-	-	**$t = 3.27$ ($p = 0.004$)**	**$t = 4.37$ ($p < 0.001$)**

In this case, the first-third principal components were used separately. Here, AUC values were averaged among 20 trials to choose tested and evaluated data set in five–fold cross-validation and their standard deviations (SD) were also derived. The paired t-value comparing with AUC only used single measure (c_1, $|c_2|$, sample entropy.) The t and corresponding p values with $p < 0.05$ are represented by bold characters. Positive t value indicates the increased AUC in the combination case.

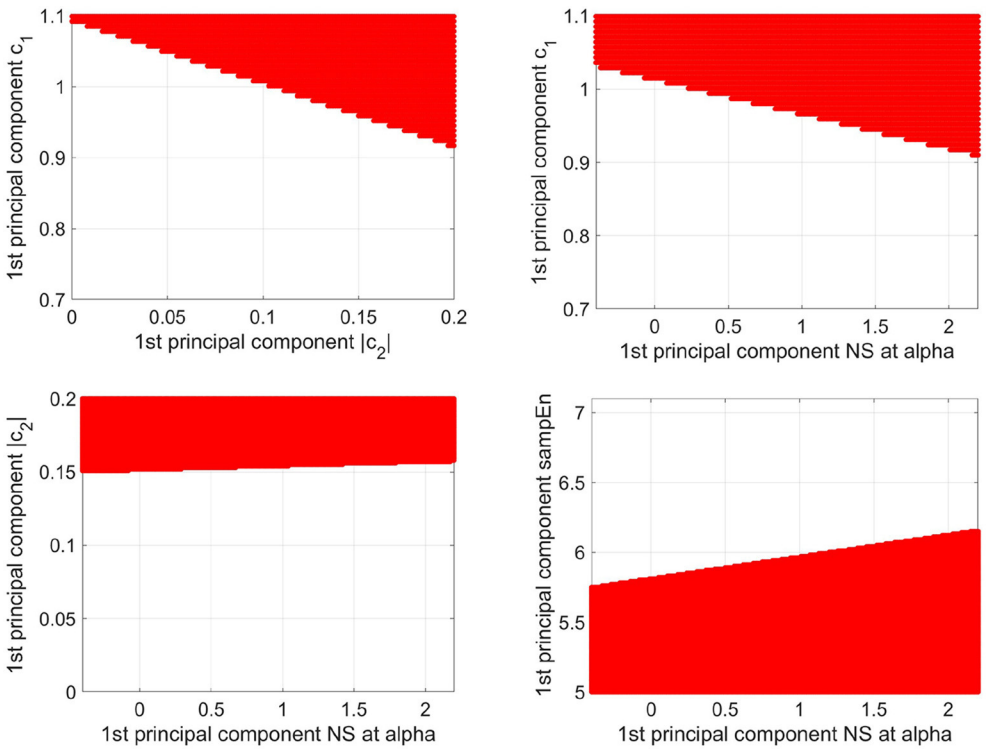


FIGURE 8 | Decision region (represented as the red region) for older participants with a decision probability of more than 0.9, obtained via logistic regression, is shown on the plane between the 1st principal component of c_1 and the 1st principal component of $|c_2|$ (upper left part), the plane between the 1st principal component of c_1 and the 1st principal component of NS at alpha (upper right part), and the plane between the 1st principal component of $|c_2|$ and the 1st principal component of NS at alpha (lower left part), and the plane between the 1st principal component of sample entropy and the 1st principal component of the NS at alpha (lower right part). Here, the other components, except for the axis of the planes, were set to average among participants in both the younger and older groups. The dependency on all these factors in the decision region was confirmed.

TABLE 7 | Younger vs. older repeated–measures ANOVA results [F value (p value, partial η^2)] with mean values of relative power at gamma band among 16 electrodes as covariate in c_1, $|c_2|$, mean values of sample entropy in low temporal-scale regions 1 to 5 (0.005–0.025 s), and NS at gamma band.

	Group	Group × node	degree of freedom (ϵ)
c_1	**$F = 9.059$ ($p = 0.004$, $\eta^2 = 0.162$)**	$F = 1.109$ ($p = 0.356$, $\eta^2 = 0.023$)	5.151 ($\epsilon = 0.343$)
$|c_2|$	**$F = 4.143$ ($p = 0.047$, $\eta^2 = 0.081$)**	$F = 1.009$ ($p = 0.420$, $\eta^2 = 0.021$)	6.023 ($\epsilon = 0.402$)
sample entropy	**$F = 10.179$ ($p = 0.003$, $\eta^2 = 0.178$)**	$F = 1.104$ ($p = 0.355$, $\eta^2 = 0.023$)	3.792 ($\epsilon = 0.253$)
NS at gamma	$F = 0.127$ ($p = 0.723$, $\eta^2 = 0.003$)	$F = 1.704$ ($p = 0.109$, $\eta^2 = 0.035$)	6.853 ($\epsilon = 0.457$)

F and p values with $p < 0.05$ are represented by bold characters. Degree of freedom and Greenhouse-Geisser adjustments ϵ in the interaction for group × node are also shown.

relationship between complexity and functional connectivity, the accuracy of aging classification improved based on the current study results.

It is imperative to discuss the reason why c_1, which is an index of smoothness, decreases with aging. Gamma activity has been reported to increase with age (Böttger et al., 2002). In our results, a higher relative power of the gamma band in the major part of the electrodes was confirmed (see **Supplementary Material**). A previous study demonstrated that the degree of complexity of EEG signals predominantly depends on smaller temporal scale (i.e., fast frequency component) behaviors, instead of on larger temporal scale (i.e., slow frequency component) behaviors (Nobukawa et al., 2019b). Our results exhibit a tendency corresponding with these findings; that is, it can be interpreted that our observed increasing complexity (decreasing c_1 and increasing smaller temporal-scale sample entropy) is induced by increasing gamma activity due to aging.

In the gamma band, artifacts due to muscle activity are larger compared to lower frequency bands (Whitham et al., 2007, 2008). Therefore, although the majority of the time segments with artifacts involving muscle activity were removed in the evaluation epochs, it was essential to investigate the influence of muscle activity on gamma band results in the estimation of functional connectivity and complexity. **Table 7** shows the results of younger vs. older groups repeated-measures ANOVA with the mean values of relative power at the gamma band among 16 electrodes as covariate in c_1, $|c_2|$, mean values of sample entropy in low temporal-scale regions 1 to 5 (0.005–0.025 s), and NS at gamma band. Resultingly, the group difference is maintained in c_1, $|c_2|$, and small-temporal-scale sample entropy. However, the group difference was not confirmed in the NS at gamma band. Therefore, in an epoch involving muscle activity, functional connectivity at the gamma band might be more strongly affected.

Moreover, we must consider the reason why the relationship between multi-fractality and complexity and their underlying neurophysiological mechanism. The time-series with large (small) multi-fractality exhibits intermittent and transient behavior with large (small) amplitude (Ihlen, 2012). Complexity reflects the degree of complexity for temporal behavior in the entire time-range, instead of intermittent behavior. Therefore, EEG signal in older subjects corresponds to homogeneous and high complex temporal behaviors. In the aging process, the connectivity of the wide range of inter neural networks becomes weak and the neural noise increases; consequently, the amount of network communication decreases (Cremer and Zeef, 1987; Onoda et al., 2012; Nobukawa et al., 2019a). Therefore, the amplitude of intermittent transient behavior driven by the neural activities from the other regions might become weak, that is, the decreased multi-fractality might reflect fewer global neural interactions. Regarding complexity, as mentioned above, increasing complexity (decreasing c_1 and increasing smaller temporal-scale sample entropy) is induced by increasing gamma activity due to aging. Considering fact that gamma-band activity relates to local excitatory and inhibitory neural interaction (Börgers and Kopell, 2003), increasing complexity is caused by the alternation of local regional neural activity.

Furthermore, it is necessary to consider why the classification accuracy is improved by adding the NS at alpha. The activity of the neural network alternates region-specifically with aging [as reviewed in Reuter-Lorenz (2002)]. In our result, NS at alpha exhibits significant high region-specificity (see **Table 3**). Such age-related region-specific characteristics could be extracted by principal component analysis and logistic regression; consequently, a relatively high classification accuracy was thus obtained in the current study. Furthermore, in recent years, studies on complexity and functional coupling have pointed out complementary relationships (Ghanbari et al., 2015; Nobukawa et al., 2020a); the studies have reported that their combination improves the detection accuracy of pathological conditions. This relationship is attributed that the inter-regional neural interactions as functional connectivities induce the local regional variability (Sporns et al., 2007; Misic et al., 2011), which was observed between NS and multi-fractality $|c_2|$ (see **Figures 6**, **7**). The complementarity was also observed in the decision plan for this study (see **Figure 8**). Based on these results, we conclude that the combination of complexity and the PLI likely improves the classification accuracy of aging.

Finally, in addition to the substantial strengths of the current investigation, the limitations of this study need to be considered. The EEG signal does not always reflect the neural activity just below the electrodes. In this study, EEG was measured using 16 electrodes, which is less than the current number of electrodes recommended by the International Federation of Clinical Neurophysiology (Seeck et al., 2017). Therefore, using MEG and high dense EEG with increased high spatial resolution and applying cortical positioning method might enhance the ability to identify the complex functional connection structures caused by aging. Regarding temporal-scale resolution, recently, Kosciessa et al. (2020) indicated an issue in the coarse-grain process's ability to rigidly extract the complexity with temporal-scale specificity (Kosciessa et al., 2020). Since the age-related alternation of power was distributed in wide frequency bands in this study, the need for a more appropriate method to extract temporal-scale dynamics is important to thoroughly investigate neural interactions and temporal-scale specific complexity. In future studies, these points should be dealt with.

5. CONCLUSION

In this study, we were able to portray the changes in neural activity with aging by using MF and MSE analyses, which are complexity analyses, as well as PLI analysis, which evaluates the functional connections. Classification accuracy was improved by combining functional connectivity, which has a complementary relationship with the index of complexity. Despite certain limitations, the outcome of this study demonstrates that the complementary relationship

between complexity and functional connectivity within EEG plays an important role in detecting age-related changes in neural activity. Therefore, these results could be useful in formulating interventions for the prevention of age-related brain dysfunction.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the informed consent did not include the declaration regarding publication of clinical data. Requests to access the datasets should be directed to Sou Nobukawa, nobukawa@cs.it-chiba.ac.jp.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Kanazawa University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MA, SN, MK, and TT designed the methods. MA and SN analyzed the results, wrote the main manuscript text, and prepared all the figures. MK conducted the experiments. All authors reviewed the manuscript.

FUNDING

This work was supported by JSPS KAKENHI for Early-Career Scientists (grant number 18K18124) (SN), Grant-in-Aid for Scientific Research (C) (grant number 20K07964) (TT), and The Okawa Foundation for Information and Telecommunications (grant number 20–20) (SN).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.793298/full#supplementary-material

REFERENCES

Aertsen, A., Gerstein, G., Habib, M., and Palm, G. (1989). Dynamics of neuronal firing correlation: modulation of "effective connectivity". *J. Neurophysiol.* 61, 900–917. doi: 10.1152/jn.1989.61.5.900

Ahmadlou, M., Adeli, H., and Adeli, A. (2011). Fractality and a wavelet-chaos-methodology for EEG-based diagnosis of Alzheimer disease. *Alzheimer Disease Assoc. Disord.* 25, 85–92. doi: 10.1097/WAD.0b013e3181ed1160

Ando, M., Nobukawa, S., Kikuchi, M., and Takahashi, T. (2021). Identification of electroencephalogram signals in Alzheimer's disease by multifractal and multiscale entropy analysis. *Front. Neurosci.* 15, 772. doi: 10.3389/fnins.2021.667614

Birren, J. E., and Fisher, L. M. (1995). Aging and speed of behavior: possible consequences for psychological functioning. *Ann. Rev. Psychol.* 46, 329–353. doi: 10.1146/annurev.ps.46.020195.001553

Börgers, C., and Kopell, N. (2003). Synchronization in networks of excitatory and inhibitory neurons with sparse, random connectivity. *Neural Comput.* 15, 509–538. doi: 10.1162/089976603321192059

Böttger, D., Herrmann, C. S., and von Cramon, D. Y. (2002). Amplitude differences of evoked alpha and gamma oscillations in two different age groups. *Int. J. Psychophysiol.* 45, 245–251. doi: 10.1016/s0167-8760(02)00031-4

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575

Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745

Costa, M., Goldberger, A. L., and Peng, C.-K. (2002). Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* 89, 068102. doi: 10.1103/PhysRevLett.89.068102

Cremer, R., and Zeef, E. J. (1987). What kind of noise increases with age? *J. Gerontol.* 42, 515–518. doi: 10.1093/geronj/42.5.515

Engels, M. M., Stam, C. J., van der Flier, W. M., Scheltens, P., de Waal, H., and van Straaten, E. C. (2015). Declining functional connectivity and changing hub locations in Alzheimer's disease: an EEG study. *BMC Neurol.* 15, 1–8. doi: 10.1186/s12883-015-0400-7

Fraschini, M., Demuru, M., Crobe, A., Marrosu, F., Stam, C. J., and Hillebrand, A. (2016). The effect of epoch length on estimated EEG functional connectivity and brain network organisation. *J. Neural Eng.* 13:036015. doi: 10.1088/1741-2560/13/3/036015

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480. doi: 10.1016/j.tics.2005.08.011

Friston, K., Frith, C., Liddle, P., and Frackowiak, R. (1993). Functional connectivity: the principal-component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* 13, 5–14. doi: 10.1038/jcbfm.1993.4

Garrett, D. D., Samanez-Larkin, G. R., MacDonald, S. W., Lindenberger, U., McIntosh, A. R., and Grady, C. L. (2013). Moment-to-moment brain signal variability: a next frontier in human brain mapping? *Neurosci. Biobehav. Rev.* 37, 610–624. doi: 10.1016/j.neubiorev.2013.02.015

Ghanbari, Y., Bloy, L., Edgar, J. C., Blaskey, L., Verma, R., and Roberts, T. P. (2015). Joint analysis of band-specific functional connectivity and signal complexity in autism. *J. Autism Develop. Disor.* 45, 444–460. doi: 10.1007/s10803-013-1915-7

Gottschalk, A., Bauer, M. S., and Whybrow, P. C. (1995). Evidence of chaotic mood variation in bipolar disorder. *Arch. Gen. Psychiatry* 52, 947–959. doi: 10.1001/archpsyc.1995.03950230061009

Hasegawa, C., Takahashi, T., Yoshimura, Y., Nobukawa, S., Ikeda, T., Saito, D. N., et al. (2018). Developmental trajectory of infant brain signal variability: a longitudinal pilot study. *Front. Neurosci.* 12:566. doi: 10.3389/fnins.2018.00566

Hata, M., Kazui, H., Tanaka, T., Ishii, R., Canuet, L., Pascual-Marqui, R. D., et al. (2016). Functional connectivity assessed by resting state EEG correlates with cognitive decline of Alzheimer's disease–an eloreta study. *Clin. Neurophysiol.* 127, 1269–1278. doi: 10.1016/j.clinph.2015.10.030

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013.05.079

Ihlen, E. A. F. E. (2012). Introduction to multifractal detrended fluctuation analysis in matlab. *Front. Physiol.* 3:141. doi: 10.3389/fphys.2012.00141

Jaffard, S., Lashermes, B., and Abry, P. (2006). "Wavelet leaders in multifractal analysis," in *Wavelet Analysis and Applications* (Basel: Springer), 201–246

Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W., and Freunberger, R. (2007). Event-related phase reorganization may explain evoked neural dynamics. *Neurosci. Biobehav. Rev.* 31, 1003–1016. doi: 10.1016/j.neubiorev.2007.03.005

Kosciessa, J. Q., Kloosterman, N. A., and Garrett, D. D. (2020). Standard multiscale entropy reflects neural dynamics at mismatched temporal scales: what's signal irregularity got to do with it? *PLoS Comput. Biol.* 16:e1007885. doi: 10.1371/journal.pcbi.1007885

Lau, T. M., Gwin, J. T., McDowell, K. G., and Ferris, D. P. (2012). Weighted phase lag index stability as an artifact resistant measure to detect

cognitive EEG activity during locomotion. *J. Neuroeng. Rehabil.* 9, 1–9. doi: 10.1186/1743-0003-9-47

McDonnell, M. D., and Ward, L. M. (2011). The benefits of noise in neural systems: bridging theory and experiment. *Nat. Rev. Neurosci.* 12, 415–425. doi: 10.1038/nrn3061

McIntosh, A. R., Kovacevic, N., and Itier, R. J. (2008). Increased brain signal variability accompanies lower behavioral variability in development. *PLoS Comput. Biol.* 4:e1000106. doi: 10.1371/journal.pcbi.1000106

Misic, B., Vakorin, V. A., Paus, T., and McIntosh, A. R. (2011). Functional embedding predicts the variability of neural activity. *Front. Syst. Neurosci.* 5, 90. doi: 10.3389/fnsys.2011.00090

Mizuno, T., Takahashi, T., Cho, R. Y., Kikuchi, M., Murata, T., Takahashi, K., et al. (2010). Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy. *Clin. Neurophysiol.* 121, 1438–1446. doi: 10.1016/j.clinph.2010.03.025

Ni, H., Zhou, L., Ning, X., Wang, L., and (ADNI), A. D. N. I. (2016). Exploring multifractal-based features for mild Alzheimer's disease classification. *Magn. Resonan. Med.* 76, 259–269. doi: 10.1002/mrm.25853

Nobukawa, S., Kikuchi, M., and Takahashi, T. (2019a). Changes in functional connectivity dynamics with aging: a dynamical phase synchronization approach. *Neuroimage* 188, 357–368. doi: 10.1016/j.neuroimage.2018.12.008

Nobukawa, S., and Nishimura, H. (2020). Synchronization of chaos in neural systems. *Front. Appl. Math. Stat.* 6, 19. doi: 10.1016/j.cnsns.2010.04.036

Nobukawa, S., Yamanishi, T., Kasakawa, S., Nishimura, H., Kikuchi, M., and Takahashi, T. (2020a). Classification methods based on complexity and synchronization of electroencephalography signals in Alzheimer's disease. *Front. Psychiatry* 11:255. doi: 10.3389/fpsyt.2020.00255

Nobukawa, S., Yamanishi, T., Nishimura, H., Wada, Y., Kikuchi, M., and Takahashi, T. (2019b). Atypical temporal-scale-specific fractal changes in Alzheimer's disease EEG and their relevance to cognitive decline. *Cogn. Neurodyn.* 13, 1–11. doi: 10.1007/s11571-018-9509-x

Nobukawa, S., Yamanishi, T., Ueno, K., Mizukami, K., Nishimura, H., and Takahashi, T. (2020b). High phase synchronization in alpha band activity in older subjects with high creativity. *Front. Human Neurosci.* 14, 420. doi: 10.3389/fnhum.2020.583049

Nolte, G., Holroyd, T., Carver, F., Coppola, R., and Hallett, M. (2004). "Localizing brain interactions from rhythmic EEG/MEG data," in *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 1, (San Francisco, CA: IEEE), 998–1001.

Nunez, P. L., Srinivasan, R., Westdorp, A. F., Wijesinghe, R. S., Tucker, D. M., Silberstein, R. B., et al. (1997). EEG coherency: I: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography Clin. Neurophysiol.* 103, 499–515. doi: 10.1016/s0013-4694(97)00066-7

Onoda, K., Ishihara, M., and Yamaguchi, S. (2012). Decreased functional connectivity by aging is associated with cognitive decline. *J. Cogn. Neurosci.* 24, 2186–2198. doi: 10.1162/jocn_a_00269

Paulus, M. P., Geyer, M. A., and Braff, D. L. (1996). Use of methods from chaos theory to quantify a fundamental dysfunction in the behavioral organization of schizophrenic patients. *Amer. J. Psychiatry* 153, 714–717. doi: 10.1176/ajp.153.5.714

Reuter-Lorenz, P. A. (2002). New visions of the aging mind and brain. *Trends Cogn. Sci.* 6, 394–400. doi: 10.1016/s1364-6613(02)01957-5

Righi, G., Tierney, A. L., Tager-Flusberg, H., and Nelson, C. A. (2014). Functional connectivity in the first year of life in infants at risk for autism spectrum disorder: an EEG study. *PLoS ONE* 9:e105176. doi: 10.1371/journal.pone.0105176

Seeck, M., Koessler, L., Bast, T., Leijten, F., Michel, C., Baumgartner, C., et al. (2017). The standardized EEG electrode array of the IFCN. *Clin. Neurophysiol.* 128, 2070–2077. doi: 10.1016/j.clinph.2017.06.254

Shim, M., Lee, S.-H., and Hwang, H.-J. (2021). Inflated prediction accuracy of neuropsychiatric biomarkers caused by data leakage in feature selection. *Sci. Rep.* 11, 1–7. doi: 10.1038/s41598-021-87157-3

Sikdar, D., Roy, R., and Mahadevappa, M. (2018). Epilepsy and seizure characterisation by multifractal analysis of EEG subbands. *Biomed. Signal Process. Control* 41, 264–270. doi: 10.1016/j.bspc.2017.12.006

Sporns, O., Honey, C. J., and Kötter, R. (2007). Identification and classification of hubs in brain networks. *PLoS One* 2:e1049. doi: 10.1371/journal.pone.0001049

Stam, C. J., Nolte, G., and Daffertshofer, A. (2007). Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum. Brain Map.* 28, 1178–1193. doi: 10.1002/hbm.20346

Stam, C. J., and Van Dijk, B. (2002). Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets. *Phys. D Nonlin. Phenomena* 163, 236–251. doi: 10.1016/S0167-2789(01)00386-4

Takahashi, T. (2013). Complexity of spontaneous brain activity in mental disorders. *Progr. Neuro Psychopharmacol. Biol. Psychiatry* 45, 258–266. doi: 10.1016/j.pnpbp.2012.05.001

Takahashi, T., Cho, R. Y., Mizuno, T., Kikuchi, M., Murata, T., Takahashi, K., et al. (2010). Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: a multiscale entropy analysis. *Neuroimage* 51, 173–182. doi: 10.1016/j.neuroimage.2010.02.009

Takahashi, T., Cho, R. Y., Murata, T., Mizuno, T., Kikuchi, M., Mizukami, K., et al. (2009). Age-related variation in EEG complexity to photic stimulation: a multiscale entropy analysis. *Clin. Neurophysiol.* 120, 476–483. doi: 10.1016/j.clinph.2008.12.043

Takahashi, T., Goto, T., Nobukawa, S., Tanaka, Y., Kikuchi, M., Higashima, M., et al. (2018). Abnormal functional connectivity of high-frequency rhythms in drug-naïve schizophrenia. *Clin. Neurophysiol.* 129, 222–231. doi: 10.1016/j.clinph.2017.11.004

Takahashi, T., Yamanishi, T., Nobukawa, S., Kasakawa, S., Yoshimura, Y., Hiraishi, H., et al. (2017). Band-specific atypical functional connectivity pattern in childhood autism spectrum disorder. *Clin. Neurophysiol.* 128, 1457–1465. doi: 10.1016/j.clinph.2017.05.010

Takahashi, T., Yoshimura, Y., Hiraishi, H., Hasegawa, C., Munesue, T., Higashida, H., et al. (2016). Enhanced brain signal variability in children with autism spectrum disorder during early childhood. *Hum. Brain Map.* 37, 1038–1050. doi: 10.1002/hbm.23089

Ueda, R., Takeichi, H., Kaga, Y., Oguri, M., Saito, Y., Nakagawa, E., et al. (2020). Atypical gamma functional connectivity pattern during light sleep in children with attention deficit hyperactivity disorder. *Brain Develop.* 42, 129–139. doi: 10.1016/j.braindev.2019.11.001

Varela, F., Lachaux, J.-P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239. doi: 10.1038/35067550

Vu, M.-A. T., Adalı, T., Ba, D., Buzsáki, G., Carlson, D., Heller, K., et al. (2018). A shared vision for machine learning in neuroscience. *J. Neurosci.* 38, 1601–1607. doi: 10.1523/JNEUROSCI.0508-17.2018

Wendt, H., and Abry, P. (2007). Multifractality tests using bootstrapped wavelet leaders. *IEEE Trans. Signal Process.* 55, 4811–4820. doi: 10.1109/TSP.2007.896269

Whitham, E. M., Lewis, T., Pope, K. J., Fitzgibbon, S. P., Clark, C. R., Loveless, S., et al. (2008). Thinking activates EMG in scalp electrical recordings. *Clin. Neurophysiol.* 119, 1166–1175. doi: 10.1016/j.clinph.2008.01.024

Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., Lewis, T., Clark, C. R., Loveless, S., et al. (2007). Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clinical neurophysiology* 118, 1877–1888. doi: 10.1016/j.clinph.2007.04.027

Yang, A. C., and Tsai, S.-J. (2013). Is mental illness complex? from behavior to brain. *Progr. Neuro Psychopharmacol. Biol. Psychiatry* 45, 253–257. doi: 10.1016/j.pnpbp.2012.09.015

Yu, M., Gouw, A. A., Hillebrand, A., Tijms, B. M., Stam, C. J., van Straaten, E. C., et al. (2016). Different functional connectivity and network topology in behavioral variant of frontotemporal dementia and Alzheimer's disease: an EEG study. *Neurobiol. Aging* 42, 150–162. doi: 10.1016/j.neurobiolaging.2016.03.018

Zorick, T., Landers, J., Leuchter, A., and Mandelkern, M. A. (2020). EEG multifractal analysis correlates with cognitive

testing scores and clinical staging in mild cognitive impairment. *J. Clin. Neurosci.* 76, 195–200. doi: 10.1016/j.jocn.2020.04.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 09 February 2022
doi: 10.3389/fnagi.2021.798410

The Effect of Aging on Brain Glucose Metabolic Connectivity Revealed by [^{18}F]FDG PET-MR and Individual Brain Networks

Nathalie Mertens[1], Stefan Sunaert[2,3], Koen Van Laere[1,4] and Michel Koole[1]*

[1] Nuclear Medicine and Molecular Imaging, Department of Imaging and Pathology, KU Leuven, Leuven, Belgium, [2] Translational MRI, Department of Imaging and Pathology, KU Leuven, Leuven, Belgium, [3] Department of Radiology, University Hospitals Leuven, Leuven, Belgium, [4] Division of Nuclear Medicine, University Hospitals Leuven, Leuven, Belgium

Contrary to group-based brain connectivity analyses, the aim of this study was to construct individual brain metabolic networks to determine age-related effects on brain metabolic connectivity. Static 40–60 min [^{18}F]FDG positron emission tomography (PET) images of 67 healthy subjects between 20 and 82 years were acquired with an integrated PET-MR system. Network nodes were defined by brain parcellation using the Schaefer atlas, while connectivity strength between two nodes was determined by comparing the distribution of PET uptake values within each node using a Kullback–Leibler divergence similarity estimation (KLSE). After constructing individual brain networks, a linear and quadratic regression analysis of metabolic connectivity strengths within- and between-networks was performed to model age-dependency. In addition, the age dependency of metrics for network integration (characteristic path length), segregation (clustering coefficient and local efficiency), and centrality (number of hubs) was assessed within the whole brain and within predefined functional subnetworks. Overall, a decrease of metabolic connectivity strength with healthy aging was found within the whole-brain network and several subnetworks except within the somatomotor, limbic, and visual network. The same decrease of metabolic connectivity was found between several networks across the whole-brain network and the functional subnetworks. In terms of network topology, a less integrated and less segregated network was observed with aging, while the distribution and the number of hubs did not change with aging, suggesting that brain metabolic networks are not reorganized during the adult lifespan. In conclusion, using an individual brain metabolic network approach, a decrease in metabolic connectivity strength was observed with healthy aging, both within the whole brain and within several predefined networks. These findings can be used in a diagnostic setting to differentiate between age-related changes in brain metabolic connectivity strength and changes caused by early development of neurodegeneration.

Keywords: metabolic connectivity, healthy aging, individual brain network, functional connectivity, [^{18}F]FDG PET

OPEN ACCESS

Edited by:
Yang Jiang,
University of Kentucky, United States

Reviewed by:
Hyejin Kang,
Seoul National University,
South Korea
Susanne Asenbaum-Nan,
Lower Landeskliniken-Holding, Austria

****Correspondence:***
Nathalie Mertens
nathalie.mertens@kuleuven.be

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *20 October 2021*
Accepted: *27 December 2021*
Published: *09 February 2022*

Citation:
Mertens N, Sunaert S, Van Laere K and Koole M (2022) The Effect of Aging on Brain Glucose Metabolic Connectivity Revealed by [^{18}F]FDG PET-MR and Individual Brain Networks. Front. Aging Neurosci. 13:798410. doi: 10.3389/fnagi.2021.798410

INTRODUCTION

[18F]FDG positron emission tomography (PET) is a valuable molecular neuroimaging technique to study the glucose metabolism in the human brain which in turn serves as a proxy for neuronal activity. Many studies have shown a progressive decrease of cerebral [18F]FDG uptake with aging, mainly observed in the medial frontal lobe and anterior cingulate cortex (Fujimoto et al., 2008; Knopman et al., 2014; Yoshizawa et al., 2014; Kakimoto et al., 2016; Ishibashi et al., 2017; Malpetti et al., 2017; Van Aalst et al., 2021). Whether these age-related changes are a linear or quadratic function of adult age with the latter showing accelerated changes in the elderly is still under debate. In parallel, structural and functional cerebral changes have been detected during the lifespan by different groups using MRI techniques. Overall, these studies showed increased gray matter (GM) atrophy observed by voxel-based morphometry MRI analyses (Good et al., 2001; Allen et al., 2005; Smith et al., 2007; Bagarinao et al., 2018) and reduced cerebral structural integrity assessed by diffusion tensor imaging (DTI) (Moseley, 2002; Head et al., 2004; Sullivan and Pfefferbaum, 2006) and differences in brain activation patterns using functional MRI (fMRI) (Grady, 2012; Avelar-Pereira et al., 2017) due to aging. To study these aging effects on brain structure and function, brain connectivity analysis has proven to be a very useful approach as it reveals important information about connections and interactions between different brain regions and allows to study the brain from a topological viewpoint. To assess brain connectivity, graph theoretical methods are generally applied which model the brain using a weighted, undirected graph. This way, a wide range of graph-based connectivity measures, reflecting both local and global brain connectivity, is extracted to quantify the underlying network topology. In literature, the majority of connectivity findings are derived from DTI and fMRI studies, which provide information on axonal pathways, or on correlations between the blood-oxygen-level-dependent (BOLD)-signal time course of different brain regions (Betzel et al., 2014; Damoiseaux, 2017; Frau-Pascual et al., 2021). In contrast to structural and functional connectivity, brain metabolic connectivity findings using [18F]FDG PET are mainly based on group-level analyses (Sala and Perani, 2019) where correlations between regional uptake values across subjects are used as connectivity measures between different brain regions. However, a novel approach using the Kullback–Leibler divergence similarity estimation (KLSE) was recently introduced to generate an individual brain metabolic network for a single subject using static [18F]FDG PET imaging (Wang et al., 2020a). This technique assumed that brain regions with similar glucose metabolism are highly interconnected while brain regions with differences in glucose metabolism have a lower connectivity strength. To determine the connectivity strength between two regions, KLSE was used to compare the intra-regional distribution of PET uptake values between different regions. Using these metabolic connectivity strengths, the approach successfully predicted the individual risk of progression from mild cognitive impairment (MCI) to Alzheimer's disease (AD) (Wang et al., 2020a). The aim of this study was to apply this novel technique on [18F]FDG PET-MR data of a cohort of 67 healthy controls, covering an age range of 20–82 years, to evaluate age-related effects on graph-based connectivity measures for network integration, segregation, and centrality. We evaluated these age effects on the level of both the whole brain and different functional brain networks where we considered networks which represent the intrinsic functional connectivity of the cerebral cortex. This study is also the first step toward using these metrics in a diagnostic setting where it is mandatory to discriminate effects of healthy aging from early development of neurodegeneration as aging is the primary risk factor for many neurodegenerative disorders (Hou et al., 2019).

MATERIALS AND METHODS

[18F]FDG PET-MR Imaging

A total of 67 healthy volunteers (33 males and 34 females; age: 52 ± 17 years, range 20–82 years) were recruited prospectively between December 2015 and February 2017. The main exclusion criteria for this study were major internal pathology or having (had) cancer, having a first-degree relative with dementia, a history of important neurological and/or psychiatric disorders, and substance abuse or current use of centrally acting medication. Subjects underwent a neurological examination resulting in a Mini-Mental State Examination (MMSE) score $\geq$ 28 and a score of $\leq$9 on the Beck's Depression Inventory (BDI) for all subjects. This study was approved by the local ethics committee of UZ Leuven Gasthuisberg, and all participants gave written informed consent.

Subjects received an intravenous bolus injection of [18F]FDG (152 ± 10 MBq) and underwent a simultaneous [18F]FDG PET-MR scan (General Electric Healthcare Signa PET-MR). Listmode data acquired between 40 and 60 min were rebinned in four frames of 5 min and corrected for motion. Sinograms were corrected for dead time, random, and scatter, while a proprietary template-based MR-based attenuation correction (MRAC) was used for attenuation correction. Each frame was reconstructed using ordered subset expectation maximization (OSEM, 28 subsets and 4 iterations) and included time of flight (TOF) information, resolution modeling, and a Gaussian post-smoothing with a full width half maximum (FWHM) of 4.5 mm. The multi-frame PET data were rescaled to standardized uptake values (SUV) and averaged to obtain a static SUV PET image.

Simultaneous with the PET acquisition, a 3D volumetric 3 Tesla T1 weighted BRAVO MR sequence was acquired using an 8-channel phased-array coil (plane: sagittal; echo time (TE): 3.2 ms; repetition time (TR): 8.5 ms; inversion time (TI): 450 ms; flip angle: 12°; receiver bandwidth: 31.2 kHz; NEX: 1; voxelsize: 1 mm × 1 mm × 1 mm) followed by a 3D T2 weighted FLAIR sequence (plane: sagittal; TE: 130 ms; TR: 8,500 ms; TI: 2,298 ms; voxelsize: 1 mm × 1 mm × 1.4 mm).

Individual Brain Metabolic Connectivity Networks

[18F]FDG PET images were spatially normalized using a non-linear normalization using the CAT12 toolbox of Statistical Parametric Mapping (SPM12; Welcome Trust Centre for Neuroimaging, University College, London, United Kingdom)

and smoothed with a Gaussian filter of 8 mm. For each subject, [18F]FDG uptake was normalized to the total uptake in the GM. Subject specific tissue probability maps for GM, white matter (WM), and cerebrospinal fluid (CSF) were derived based on the 3D T1-weighted MR data in SPM12 and used to delineate volume of interest (VOI) defined by the Schaefer atlas (Schaefer et al., 2018). This functional atlas includes the frontoparietal (4 VOIs) network together with seven functional subnetworks containing the visual (13 VOIs), somatomotor (14 VOIs), dorsal attention (13 VOIs), salience and ventral attention (14 VOIs), limbic (5 VOIs), control (16 VOIs), and default mode (21 VOIs) network (**Supplementary Table 1**).

For depicting an individual metabolic network for each subject, the 100 brain parcels, determined by the Schaefer atlas, were considered as nodes, and the [18F]FDG uptake in each pair of nodes was used to generate a metabolic correlation matrix. This was performed by extracting the intensity values of voxels within each node to estimate the probability density function (PDF) of intensity values for that parcel. All PDFs were estimated using brain parcels in MNI space containing minimum 800 voxels. Furthermore, the PDF for each parcel was estimated using the kernel density estimation (KDE) with optimal bandwidths for the number of voxels of that specific parcel chosen automatically using the diffusion Botev method as implemented in the KDE-diffusion toolbox in Python version 3.9 (Botev et al., 2010). In addition, PDFs were estimated in a standardized histogram space with a fixed range of values and a fixed bin size for all parcels and all subjects. Then, the KLSE method was used to estimate the similarity between the PDFs of two nodes and construct a correlation matrix which represents the pairwise metabolic connections or edges. In general, KLSE is based on the KL divergence (D_{KL}) between two PDFs. However, to have a symmetric measure, the following variation of D_{KL} was used:

$$D_{KL}(P, Q) = \int_x \left(P(x)\ log\frac{P(x)}{Q(x)} + Q(x)\ log\frac{Q(x)}{P(x)} \right) dx$$

where P and Q are two PDFs defined on the same x range. Finally, the metabolic connectivity strength between two nodes was calculated as the KL similarity (KLS) measure as follows:

$$KLS(P, Q) = e^{-D_{KL}(P,Q)}.$$

This way, an undirected weighted metabolic connectivity matrix was estimated for each subject and quantified using graph-based connectivity metrics without applying a threshold to generate a binarized connectivity matrix.

Brain Metabolic Connectivity Metrics

Several graph theory metrics of metabolic connectivity were calculated to characterize global and nodal connectivity. Global connectivity of each network was assessed using the mean connectivity strength and the characteristic path length, while nodal connectivity was assessed using the clustering coefficient and the local efficiency. All nodal metrics were averaged over all pairs of nodes in order to examine network characterization of the whole network. Furthermore, to assess network centrality, four nodal metrics were used, being the degree, characteristic path length, clustering coefficient, and betweenness centrality to determine central nodes within the network, called hubs. To calculate the clustering coefficient and the local efficiency, a generalization for weighted undirected graphs was used as proposed by Wang et al. (2017), while all other metrics were calculated using the brain-connectivity toolbox in Python (Rubinov and Sporns, 2010).

First, to quantify the connectivity within each individual network, the average metabolic connectivity strength over all pairs of nodes of each network was determined for the whole-brain network. In addition to this whole-brain connectivity measure, the metabolic connectivity strength within each functional subnetwork was assessed as well as between-network metabolic connectivity strengths. The average within-network connectivity strength was calculated by averaging the connectivity values over all nodes within each functional subnetwork, while between-network connectivity strengths were generated by averaging the connectivity values over all nodes within two functional subnetworks (Varangis et al., 2019).

Then, two types of connectivity metrics were calculated for each metabolic network. First, the characteristic path length of each network was calculated as the measure of functional integration of the brain network. Second, the average clustering coefficient over all nodes, reflecting the average prevalence of clustered connectivity around individual nodes, as well as the average local efficiency which represents the average strength of local connectedness within neighboring nodes, were calculated as measures of functional segregation of the brain network. Measures of network integration and segregation were calculated for the whole-brain network, as well as within each functional subnetwork.

Finally, hubs within each individual network were identified based on the hub score using four criteria which are determined based on whether the node belongs to the top 20% of nodes (a) showing the highest degree, (b) showing the lowest path length, (c) showing the lowest clustering coefficient, and (d) showing the highest betweenness centrality. If the hub score was at least 2, the node was considered a hub (Ran et al., 2020). As such, the number of hubs is a measure of functional centrality in the corresponding network. To examine a potential network reorganization during aging, we divided our study population in a group of young ($n = 22$, age: 32 ± 7 years), middle-aged ($n = 22$, age: 51 ± 5 years), and elderly ($n = 23$, age: 69 ± 6 years) healthy volunteers and compared the number of hubs within the whole brain and within functional subnetworks between these groups.

Statistics

A multiple linear regression model was used to assess the effect of age on the different network connectivity metrics within the whole-brain network, as well as within and between different functional subnetworks. Both a linear and second-order polynomial (quadratic) age dependency of connectivity metrics were considered while the sum-of-squares F test was used to select the most appropriate model. Goodness of fit was reported using the coefficient of multiple correlation r. All statistical analyses were performed with Prism (version 9, GraphPad, San Diego, CA, United States) using a significance level of $p < 0.05$.

Then, significant multiple linear regression models were used to assess% differences in connectivity metrics during the adult lifespan by comparing a 20-year-old subject with an 80-year-old subject.

RESULTS

An overview of linear and quadratic age effects on metabolic connectivity metrics within the whole-brain network and the functional subnetworks are given in **Table 1**. In addition, representative mean and coefficient of covariation of metabolic connectivity matrices for three age groups (young, middle-aged and old, respectively) are shown in **Supplementary Figure 1**.

Age Effects on Mean Metabolic Connectivity Strength

An overview of regression analysis results assessing the average metabolic connectivity strength as function of age is given in **Table 2** and **Figure 1** for the whole-brain network as well as for different functional subnetworks. Within the whole network, a linear decreasing age effect on the metabolic connectivity strength was found ($p = 0.0001$, $r = 0.45$), resulting in a decrease of 16.3% during the adult lifespan. This decrease in metabolic strength within the brain network was also observed when comparing the distribution of metabolic strength of a 20-year-old with an 80-year-old subject (**Supplementary Figure 2**). For the predefined functional subnetworks, a quadratic decreasing age effect on the metabolic connectivity strength was found within the frontoparietal network ($p = 0.0008$, $r = 0.46$), showing a decrease in metabolic strength of 41.2% between a 20-year-old and 80-year-old subject. In the default mode, control, dorsal attention, and ventral attention network, a linear decrease of metabolic connectivity strength with age was found ($p = 0.0179$, $r = 0.29$; $p = 0.0048$, $r = 0.34$; $p = 0.0061$, $r = 0.33$; and $p < 0.0001$, $r = 0.48$, respectively), resulting in a decrease of 14.9, 17.2, 18.4, and 28.3%, respectively, during the adult lifespan. In contrast, no effect of age was found in the somatomotor, limbic, and visual network. Representative connectome networks of the ventral attention and the somatomotor network with an upper connectivity threshold of 0.80 of a young and an old healthy subject are given in **Figure 2**, showing lower connectivity within the older subject compared with the younger subject in the ventral attention, dorsal attention, frontoparietal, control, and default mode network, but not in the somatomotor, limbic, and visual network.

Finally, 26 out of 36 (72%) between-network metabolic connectivity strengths showed a significant decrease with age (**Supplementary Table 2**). In 7 out of 36 (19%) between-network connectivity strengths, the quadratic model was the preferred model to model age effects, while 19 out of 36 (53%) between-networks showed a linear decrease with age.

Age Effects on Functional Integration Metrics

Results of the regression analyses to model a functional integration metric assessed by the characteristic path length as a function of age within the whole-brain network and within functional subnetworks are given in **Table 2** and **Figure 3**. For the characteristic path length, a linear increasing effect of age was found within the whole-brain network ($p < 0.001$, $r = 0.67$), resulting in an increase of 13.2% during the adult lifespan. Within functional subnetworks, a quadratic increasing effect of age was found within the frontoparietal network ($p = 0.0001$, $r = 0.49$). During the adult lifespan, an increase in a characteristic path length of 51.4% was found within this network. Otherwise, in the default mode, control, dorsal attention, and ventral attention network, a linear increase of network integration was found with age as assessed by the characteristic path length ($p = 0.0036$, $r = 0.35$; $p = 0.0121$, $r = 0.30$; $p = 0.0031$, $r = 0.36$; and $p < 0.0001$, $r = 0.51$, respectively). Comparing the characteristic path length of a 20-year-old and 80-year-old subject, an increase of 16.8, 13.9, 20.8, and 34.4% was found in the default mode, control, dorsal attention, and ventral attention network, respectively. In contrast, no effect of age on the characteristic path length was found in the somatomotor, limbic, and visual network.

Age Effects on Functional Segregation Metrics

Results of the regression analyses to model functional segregation metrics assessed by the average clustering coefficient and the average local efficiency, as function of age within the whole-brain network and within functional subnetworks, are given in **Table 2** and **Figures 4**, **5**. For both the average clustering coefficient

TABLE 1 | Overview of linear and quadratic age effects (age and age^2, respectively) on metabolic network characteristics within the whole-brain network, as well as within functional subnetworks obtained from a multiple linear regression model.

	Mean connectivity strength	Characteristic path length	Average clustering coefficient	Average local efficiency
Whole brain network	Age	Age	Age	Age
Frontoparietal network	Age^2	Age^2	Age^2	Age^2
Default mode network	Age	Age	Age	Age^2
Control network	Age	Age	Age	Age
Dorsal attention network	Age	Age	Age	Age
Ventral attention network	Age	Age	Age	Age
Somatomotor network	/	/	/	/
Limbic network	/	/	/	/
Visual network	/	/	/	/

TABLE 2 | Overview of multiple linear regression analyses to model network metrics as function of age within the whole-brain network and within functional subnetworks.

	p-Value	*r*-Value	ß0	ß1	ß2	20y	80y	%Diff
Mean connectivity strength								
Whole brain network	**0.0001**	0.45	**0.34**	**-0.88E-03**	/	0.32	0.27	–16.3
Frontoparietal network	**0.0003**	0.48	0.13	**6.80E-03**	**–0.08E-03**	0.23	0.14	–41.2
Default mode network	**0.0179**	0.29	**0.31**	**-0.739E-03**	/	0.30	0.25	–14.9
Control network	**0.0048**	0.34	**0.44**	**–1.182E-03**	/	0.41	0.34	–17.2
Dorsal attention network	**0.0061**	0.33	**0.50**	**–1.442E-03**	/	0.47	0.38	–18.4
Ventral attention network	**<0.0001**	0.48	**0.48**	**–2.06E-03**	/	0.44	0.31	–28.3
Somatomotor network	0.9922	/	/	/	/	/	/	/
Limbic network	0.4787	/	/	/	/	/	/	/
Visual network	0.2913	/	/	/	/	/	/	/
Characteristic path length								
Whole brain network	**<0.0001**	0.49	**2.29**	**5.26E-03**	/	2.39	2.71	13.2
Frontoparietal network	**0.0001**	0.49	**4.56**	**–94.12E-03**	**1.21E-03**	3.16	4.78	51.4
Default mode network	**0.0036**	0.35	**2.43**	**7.19E-03**	/	2.57	3.01	16.8
Control network	**0.0121**	0.30	**1.97**	**4.78E-03**	/	2.07	2.35	13.9
Dorsal attention network	**0.0031**	0.36	**1.69**	**6.28E-03**	/	1.81	2.19	20.8
Ventral attention network	**<0.0001**	0.51	**1.72**	**11.11E-03**	/	1.94	2.61	34.4
Somatomotor network	0.9621	/	/	/	/	/	/	/
Limbic network	0.5648	/	/	/	/	/	/	/
Visual network	0.4884	/	/	/	/	/	/	/
Average clustering coefficient								
Whole brain network	**0.0001**	0.45	**0.36**	**–0.92E-03**	/	0.34	0.28	–16.4
Frontoparietal network	**0.0003**	0.47	**0.23**	**9.40E-03**	**–0.11E-03**	0.37	0.28	–24.8
Default mode network	**0.0314**	0.26	**0.34**	**–0.74E-03**	/	0.33	0.29	–13.5
Control network	**0.0048**	0.34	**0.49**	**–1.31E-03**	/	0.46	0.39	–16.9
Dorsal attention network	**0.0117**	0.31	**0.58**	**–1.54E-03**	/	0.55	0.45	–16.9
Ventral attention network	**<0.0001**	0.47	**0.55**	**–2.29E-03**	/	0.51	0.37	–27.0
Somatomotor network	0.9839	/	/	/	/	/	/	/
Limbic network	0.9913	/	/	/	/	/	/	/
Visual network	0.2458	/	/	/	/	/	/	/
Average local efficiency								
Whole brain network	**<0.0001**	0.50	**0.24**	**–0.60E-03**	/	0.23	0.20	–15.5
Frontoparietal network	**0.0011**	0.44	0.05	**8.31E-03**	**–0.10E-03**	0.18	0.08	–53.4
Default mode network	**0.0134**	0.36	**0.15**	2.22E-03	**–0.03E-03**	0.18	0.15	–16.1
Control network	**0.0026**	0.36	**0.29**	**–0.93E-03**	/	0.27	0.22	–20.4
Dorsal attention network	**0.0096**	0.31	**0.34**	**–1.04E-03**	/	0.32	0.26	–19.5
Ventral attention network	**0.0001**	0.46	**0.29**	**–1.36E-03**	/	0.27	0.19	–30.5
Somatomotor network	0.9425	/	/	/	/	/	/	/
Limbic network	0.1944	/	/	/	/	/	/	/
Visual network	0.8200	/	/	/	/	/	/	/

Multiple linear regressions are described as $Y = ß0 + ß1.age + ß2.age^2$. Regression p-values, overall r-values, and regression coefficients are given. Significant p-values and coefficients are given in bold. Metric values and differences between an 80-year-old and 20-year-old subject are also given.

and the average local efficiency, a linear decreasing effect of age was found within the whole-brain network ($p = 0.0001$, $r = 0.45$ and $p < 0.0001$, $r = 0.50$). This resulted in a decrease of 16.4 and 15.5%, respectively, in terms of average clustering coefficient and average local efficiency during the adult lifespan. For the functional subnetworks, a decreasing quadratic effect with age was observed in terms of average local efficiency within the frontoparietal network ($p = 0.0011$, $r = 0.44$), as well as within the default mode network ($p = 0.0134$, $r = 0.36$). Within these networks, a decrease in the average local efficiency of 53.4 and 16.1% was observed between a 20-year-old and 80-year-old subject. For the average clustering coefficient, a decreasing quadratic effect with age was also found within the frontoparietal network ($p = 0.0003$, $r = 0.47$), while a decreasing linear effect with age was observed within the default network ($p = 0.0314$, $r = 0.26$). During the adult lifespan, a decrease in the average clustering coefficient of 24.8 and 13.5%, respectively, was observed within these networks. Furthermore, a decreasing linear age effect of network segregation was also found in the control ($p = 0.0048$, $r = 0.34$), dorsal attention ($p = 0.0117$, $r = 0.31$), and ventral attention network ($p < 0.0001$, $r = 0.47$) as assessed by the average clustering coefficient. For these subnetworks, the

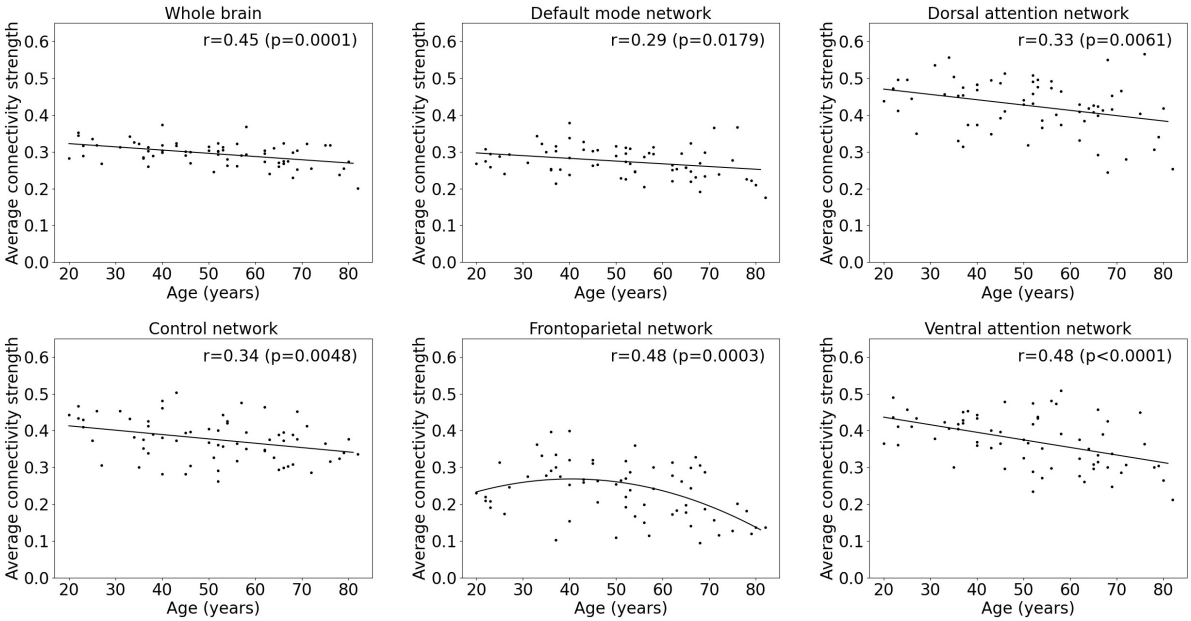


FIGURE 1 | Multiple linear regression model of average metabolic connectivity strength with age within the whole-brain network and within functional subnetworks.

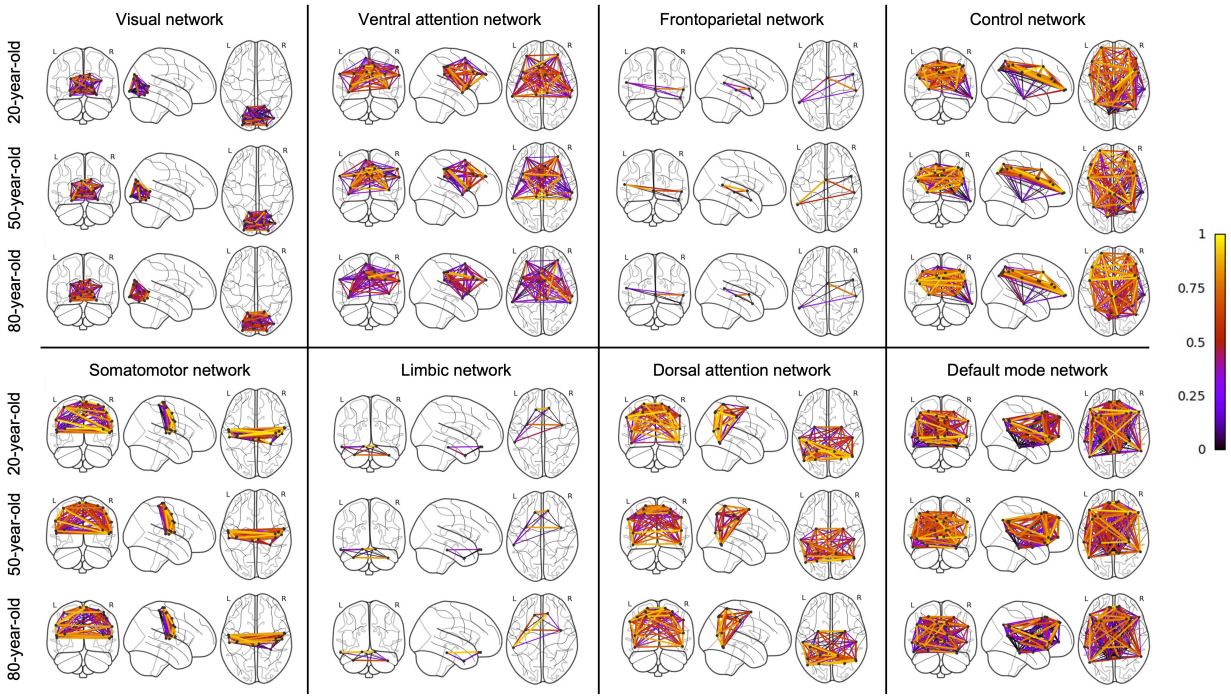


FIGURE 2 | Connectome for a young and elderly healthy subject within the ventral attention network and the somatomotor network with an upper threshold of 0.80 for the metabolic connectivity strength, showing a decreased metabolic connectivity strength with age in the ventral attention network but not in the somatomotor network.

average local efficiency also showed a decreasing linear effect with age (p = 0.0026, r = 0.36; p = 0.0096, r = 0.31; and p = 0.0001, r = 0.46). Comparing a 20-year-old subject with an 80-year-old subject showed a decrease in the average clustering coefficient of 16.9%, 16.9%, and 27.0% respectively, and a decrease in the average local efficiency of 20.4, 19.5, and 30.5%, respectively. Again, no effect of age was found in the somatomotor, limbic, and visual network.

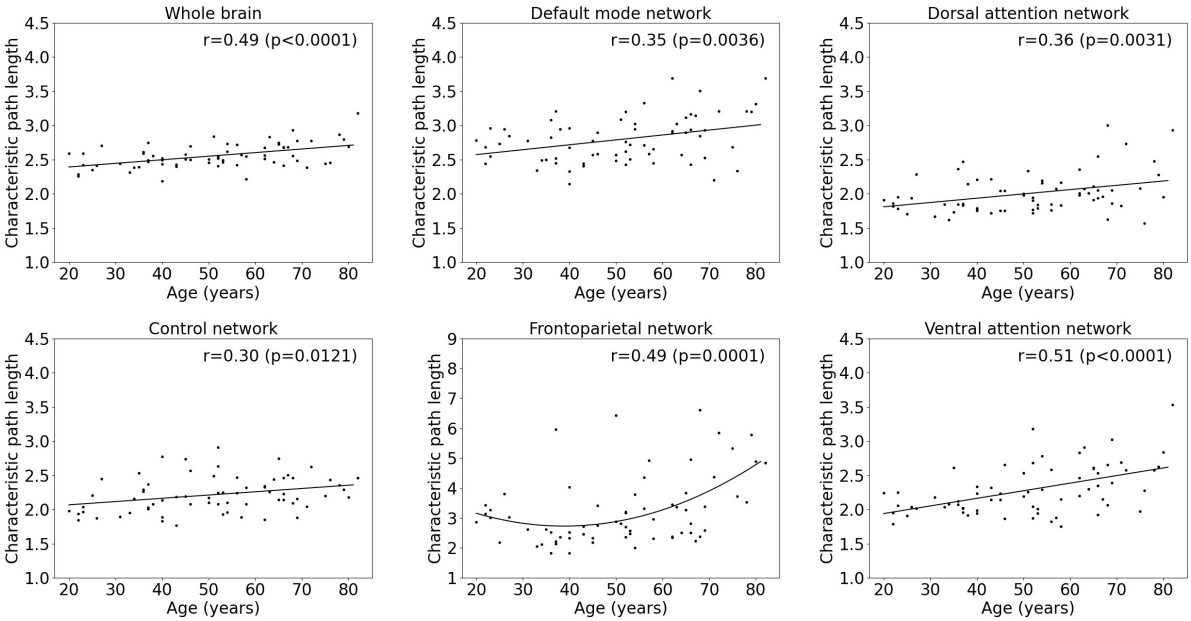


FIGURE 3 | Multiple linear regression model of the characteristic path length with age within the whole-brain network and within functional subnetworks.

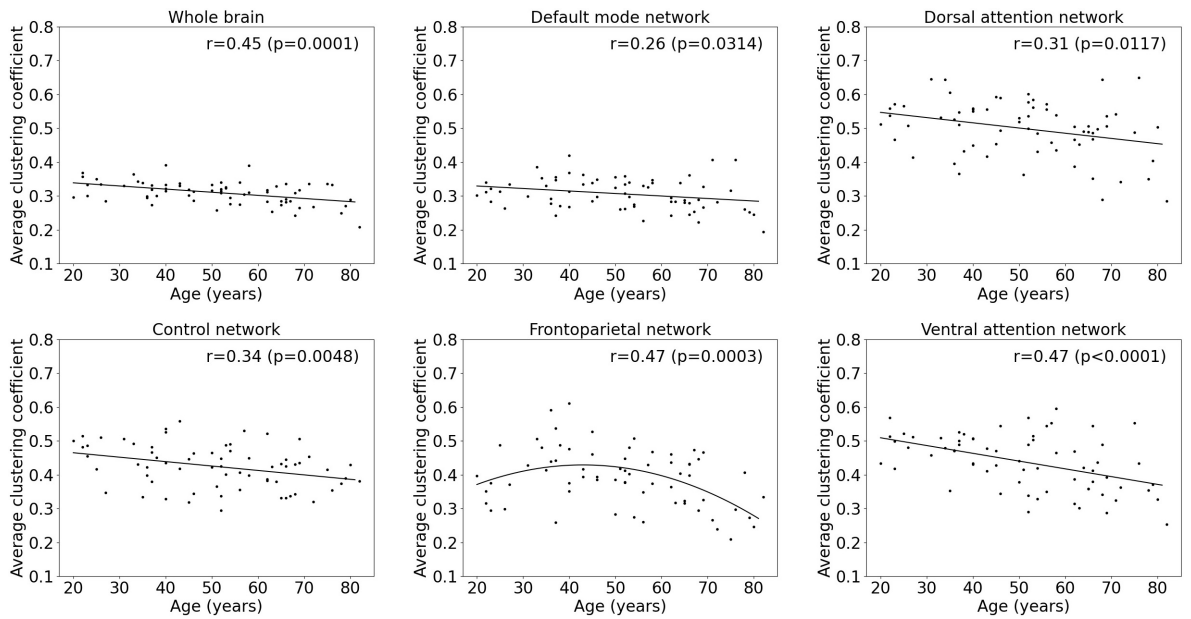


FIGURE 4 | Multiple linear regression model of the average clustering coefficient with age within the whole-brain network and within functional subnetworks.

Age Effects on Functional Centrality Metrics

The median and interquartile range (IQR) of the absolute number of hubs within the whole brain as well as within the functional subnetworks is given in **Table 3**. Within the whole brain, a linear regression analysis did not show an effect of age on the number of hubs (p = 0.20). In addition, no age-related reorganization was observed as the number of hubs within the whole brain and functional subnetworks remained stable in a young, middle-aged, and old group (**Table 3**).

DISCUSSION

To our knowledge, this is the first study to explore the effect of age on individual brain metabolic connectivity using [18F]FDG PET. Our results showed that individual brain metabolic networks

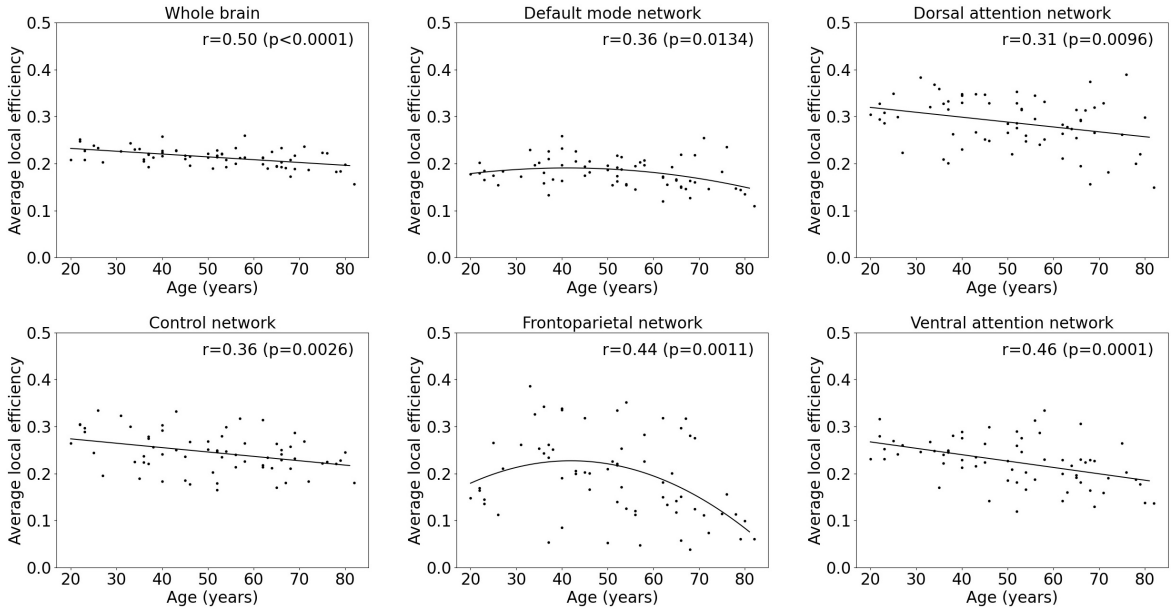


FIGURE 5 | Multiple linear regression model of the average local efficiency with age within the whole-brain network and within functional subnetworks.

became less integrated and less segregated during aging. We observed these age-related effects on the level of the whole brain as well as within the functional subnetworks except within the control, limbic, and visual network. The same decrease of metabolic connectivity was found between several networks across the whole-brain network and the functional subnetworks. Meanwhile, no network reorganization was observed with aging as distribution of hubs throughout the brain, and different subnetworks remained unchanged during aging.

In literature, only few studies explored the effect of aging on metabolic connectivity networks in healthy subjects. While our study found a decrease in metabolic connectivity with aging, two other groups reported an opposite trend using a group-based correlation approach (Arnemann et al., 2018; Huang et al., 2021). However, it is not straightforward to compare our findings with these studies as a group-based correlation approach highly depends on group composition where a homogeneous group with low inter-subject variability in regional [18F]FDG uptake could result in lower correlation measures. In addition, correlation measures are not sensitive to differences in regional uptake values which are consistent across subjects. On the contrary, the KLSE approach compares the intra-regional metabolic distribution between different regions within a single subject where it uses all uptake values within each brain region and thus much more information compared with correlation-based measures which consider only the averaged regional uptake of brain regions. As such, this approach provides a quantitative representation of the tracer distribution throughout the brain and the different subnetworks with a high average metabolic strength between nodes representing a rather homogeneous [18F]FDG uptake in the corresponding brain regions. This way, metabolic connectivity estimated across subjects using correlation measures and with-subject using the KLSE approach proved to be high complementary. Meanwhile, univariate VOI-based and voxelwise approaches have mainly been used so far to study age-related effects on brain glucose metabolism. These studies showed differences in the age-dependency of glucose metabolism between different brain regions with consistently higher age-related decrease of [18F]FDG uptake in the frontal, cingulate, temporal cortex, and in subcortical GM regions compared with other brain regions (Fujimoto et al., 2008; Knopman et al., 2014; Yoshizawa et al., 2014; Kakimoto et al., 2016; Ishibashi et al., 2017; Malpetti et al., 2017; Van Aalst et al., 2021). As this translates into a more heterogeneous distribution of [18F]FDG uptake throughout the brain with aging, this also results in a lower global metabolic connectivity strength in elderly healthy persons which is in line with our findings. More specifically, these VOI-based results, showing higher age-related decrease of [18F]FDG uptake in prefrontal cortex, medial frontal cortex, temporal and high

TABLE 3 | Overview of number of hubs within the whole-brain network, as well as within functional subnetworks.

	All subjects	Young	Middle-aged	Old
Whole brain network	19 (18–21)	20 (18–21)	19 (18–21)	20 (19–22)
Default mode network	4 (3–5)	4 (3–4)	4 (3–5)	4 (3–5)
Control network	4 (3–5)	4 (3–4)	3 (2–4)	4 (3–5)
Dorsal attention network	3 (2–4)	4 (3–5)	3 (2–4)	3 (3–4)
Somatomotor network	3 (2–4)	2 (1–3)	3 (2–4)	3 (3–5)
Ventral attention network	2 (1–3)	2 (1–3)	2 (1–3)	2 (1–3)
Visual network	1 (0–2)	1 (0–2)	1 (0–2)	1 (0–2)
Limbic network	1 (0–2)	1 (1–2)	1 (0–2)	1 (0–2)
Frontoparietal network	1 (0–1)	1 (0–2)	1 (1–2)	1 (0–1)

Values are presented as median [interquartile range (IQR)] of all subjects, as well as of a young, middle-aged, and old group.

parietal cortex, and insula (Van Aalst et al., 2021), also support our findings of an age-related decrease of metabolic connectivity in default mode, frontoparietal, control, dorsal attention, and ventral attention network as these brain regions are involved in these functional networks.

Contrary to metabolic connectivity, the effect of age on functional connectivity has been extensively explored using fMRI (Betzel et al., 2014; Geerligs et al., 2014; Damoiseaux, 2017; Farras-Permanyer et al., 2019; Varangis et al., 2019; Mancho-Fora et al., 2020). In general, functional connectivity based on fMRI showed an age-related decrease in functional connectivity and a loss in network integrity as well as in network segregation during the lifespan, which are in line with our metabolic connectivity findings. Within subnetworks, Varangis et al. (2019) reported decreased functional connectivity based on fMRI in the default mode, frontoparietal, ventral attention, and dorsal attention network which is similar as our age-related changes within the functional subnetworks using [18F]FDG PET. To compare our findings with these fMRI studies on functional connectivity, we used the Schaefer atlas for the brain parcellation as this allowed us to define the major functional networks across the cerebral cortex which are also frequently used in fMRI analysis (Watabe and Hatazawa, 2019). Although connectivity metrics are not directly comparable, we observed a similar decrease in local efficiency and increase in characteristic path length and clustering coefficient (Sun et al., 2012).

In terms of methodology, we applied a novel KLSE approach to define metabolic connectivity at a subject level. This approach has already been validated and successfully implemented to predict the progression from MCI to AD using [18F]FDG PET (Wang et al., 2020a,b). To determine an individual metabolic network, the KLSE approach relies on a predefined atlas for brain parcellation and defining the nodes of each individual network. To assure a robust estimation of the PDF within each parcel, a granularity of 100 parcels was selected for the whole-brain parcellation as more parcels would result in less voxels per parcel and impact PDF estimates (Brownlee, 2020). In addition, a fully weighted network approach was chosen to preserve the higher information content over binary network and avoid the need for a rather arbitrary threshold for the binarization. Although weighted networks are often more difficult to interpret, weighted networks are especially of interest for studying brain metabolic connectivity as variations in metabolic connectivity strength can be described by connectivity weights (Fornito et al., 2016). Finally, we also implemented the KLSE approach including a region-based voxelwise correction for partial volume effects (PVC) to assess the contribution of underlying morphology changes on aging (Thomas et al., 2011; Greve et al., 2016). Age-related results from metabolic connectivity based on PVC [18F]FDG PET images also showed a less integrated and less segregated metabolic brain network during aging within the whole brain and the same functional subnetworks (**Supplementary Table 3**). In 18 out of 36 (50%) between-networks, a similar age-related decrease in metabolic connectivity was found across the whole brain and several functional subnetworks (**Supplementary Table 4**). Altogether, age-related results from metabolic connectivity based on PVC [18F]FDG PET images agreed with the uncorrected PET FDG results (**Supplementary Table 3**), suggesting a true observed effect of aging on the metabolic connectivity.

In terms of study limitations, we considered only a limited number of connectivity metrics. However, we made sure to include metrics that represented both global and local brain metabolic connectivity and measured network integration, segregation, and centrality such that a wide range of connectivity metrics was covered. Another limitation is that we did not look at a gender effect on the metabolic connectivity due to rather small sample size and because man and women were not homogeneously distributed within our study population.

In the future, it would be interesting to further explore this novel individual approach for PET tracers targeting specific neurotransmission systems, amyloid load, or tau deposition (Sala and Perani, 2019). For these tracers, single-subject network metrics could serve as diagnostic markers to quantify differences between healthy subjects and specific patient groups and to explore the association between these individual metrics and clinical outcome (Paldino et al., 2017; Fortier et al., 2019). However, a diagnostic approach would probably benefit much more from using a brain atlas for the parcellation, which is more related to PET data analysis, such as the Hammers atlas, instead of the Schaefer atlas which is more functionally oriented. Furthermore, one could take advantage of this individual approach and combine this individual metabolic network with individual structural and functional networks to obtain an integrated multiplex network. This way, the network topology of the human brain can be explored from a multilayer perspective which could further improve diagnosis and patient stratification (Giuliano Zippo and Castiglioni, 2016).

CONCLUSION

In this study, age-related changes in brain metabolic connectivity during the adult lifespan were revealed using an individual brain metabolic network constructed with [18F]FDG PET. Overall, metabolic connectivity within the whole-brain network decreased with aging and resulted into a less integrated and less segregated network, while no evidence was found for reorganization of brain metabolic networks during healthy aging. The same age-related decrease in metabolic connectivity was also found in predefined functional subnetworks but not in the control, limbic, and visual network. A similar decrease in metabolic connectivity with aging was observed between several networks across the whole brain and functional subnetworks. Finally, these findings were in line with age-related functional connectivity changes during the adult lifespan using fMRI.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are not publicly available due to privacy/ethical restrictions. Upon reasonable request, the anonymized data could be shared on approval by the local ethics committee. Requests to access the datasets should be directed to NM, nathalie.mertens@kuleuven.be.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Local Ethics Committee of UZ Leuven Gasthuisberg. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

NM and MK contributed to the analysis of data and design of this study. NM, SS, KVL, and MK contributed to interpretation of data and design of this study. All authors critically revised the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge the contributions by Stefanie Willekens to data acquisition, Kim Serdons and the radiopharmacy team to tracer production, Kwinten Porters, Jef Van Loock, Guido Putzeys, Kris Byloos, and Stefan Ghysels to data acquisition, and Ronald Peeters from the PET/MR physics team to general support of the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2021.798410/full#supplementary-material

REFERENCES

Allen, J. S., Bruss, J., Brown, C. K., and Damasio, H. (2005). Normal neuroanatomical variation due to age: the major lobes and a parcellation of the temporal region. *Neurobiol. Aging* 26, 1245–1260. doi: 10.1016/j.neurobiolaging.2005.05.023

Arnemann, K. L., Stöber, F., Narayan, S., Rabinovici, G. D., and Jagust, W. J. (2018). Metabolic brain networks in aging and preclinical Alzheimer's disease. *Neuroimage Clin.* 17, 987–999. doi: 10.1016/j.nicl.2017.12.037

Avelar-Pereira, B., Bäckman, L., Wåhlin, A., Nyberg, L., and Salami, A. (2017). Age-related differences in dynamic interactions among default mode, frontoparietal control, and dorsal attention networks during resting-state and interference resolution. *Front. Aging Neurosci.* 9:152. doi: 10.3389/fnagi.2017.00152

Bagarinao, E., Watanabe, H., Maesawa, S., Mori, D., Hara, K., Kawabata, K., et al. (2018). An unbiased data-driven age-related structural brain parcellation for the identification of intrinsic brain volume changes over the adult lifespan. *Neuroimage* 169, 134–144. doi: 10.1016/j.neuroimage.2017.12.014

Betzel, R. F., Byrge, L., He, Y., Goñi, J., Zuo, X. N., and Sporns, O. (2014). Changes in structural and functional connectivity among resting-state networks across the human lifespan. *Neuroimage* 102, 345–357. doi: 10.1016/j.neuroimage.2014.07.067

Botev, Z., Grotowski, J., and Kroese, D. (2010). Kernel density estimation via diffusion. *Ann. Stat.* 38, 2916–2957.

Brownlee, J. (2020). *Probability for Machine Learning*. Amsterdam: Elsevier.

Damoiseaux, J. S. (2017). Effects of aging on functional and structural brain connectivity. *Neuroimage* 160, 32–40. doi: 10.1016/j.neuroimage.2017.01.077

Farras-Permanyer, L., Mancho-Fora, N., Montalà-Flaquer, M., Bartrés-Faz, D., Vaqué-Alcázar, L., Peró-Cebollero, M., et al. (2019). Age-related changes in resting-state functional connectivity in older adults. *Neural Regen .Res.* 14, 1544–1555. doi: 10.4103/1673-5374.255976

Fornito, A., Zalesky, A., and Bullmore, E. (2016). . *Fundamentals of Brain Network Analysis*. Cambridge, MA: Academic Press.

Fortier, M., Castellano, C.-A., Croteau, E., Langlois, F., Bocti, C., St-Pierre, V., et al. (2019). A ketogenic drink improves brain energy and some measures of cognition in mild cognitive impairment. *Alzheimers Dement.* 15, 625–634. doi: 10.1016/j.jalz.2018.12.017

Frau-Pascual, A., Augustinack, J., Varadarajan, D., Yendiki, A., Salat, D. H., Fischl, B., et al. (2021). Conductance-based structural brain connectivity in aging and dementia. *Brain Connect.* 11, 566–583. doi: 10.1089/brain.2020.0903

Fujimoto, T., Matsumoto, T., Fujita, S., Takeuchi, K., Nakamura, K., Mitsuyama, Y., et al. (2008). Changes in glucose metabolism due to aging and gender-related differences in the healthy human brain. *Psychiatry Res. Neuroimaging* 164, 58–72. doi: 10.1016/j.pscychresns.2006.12.014

Geerligs, L., Renken, R. J., Saliasi, E., Maurits, N. M., and Lorist, M. M. (2014). A brain-wide study of age-related changes in functional connectivity. *Cereb. Cortex* 25, 1987–1999. doi: 10.1093/cercor/bhu012

Giuliano Zippo, A., and Castiglioni, I. (2016). Integration of 18FDG-PET metabolic and functional connectomes in the early diagnosis and prognosis of the Alzheimer's disease. *Curr. Alzheimer Res.* 13, 487–497. doi: 10.2174/1567205013666151116142451

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., and Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* 14, 21–36. doi: 10.1006/nimg.2001.0786

Grady, C. (2012). The cognitive neuroscience of ageing. *Nat. Rev. Neurosci.* 13, 491–505.

Greve, D. N., Salat, D. H., Bowen, S. L., Izquierdo-Garcia, D., Schultz, A. P., Catana, C., et al. (2016). Different partial volume correction methods lead to different conclusions: an 18F-FDG-PET study of aging. *Neuroimage* 132, 334–343. doi: 10.1016/j.neuroimage.2016.02.042

Head, D., Buckner, R. L., Shimony, J. S., Williams, L. E., Akbudak, E., Conturo, T. E., et al. (2004). Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the alzheimer type: evidence from diffusion tensor imaging. *Cereb. Cortex* 14, 410–423. doi: 10.1093/cercor/bhh003

Hou, Y., Dan, X., Babbar, M., Wei, Y., Hasselbalch, S. G., Croteau, D. L., et al. (2019). Ageing as a risk factor for neurodegenerative disease. *Nat. Rev. Neurol.* 15, 565–581. doi: 10.1038/s41582-019-0244-7

Huang, Q., Ren, S., Zhang, T., Li, J., Jiang, D., Xiao, J., et al. (2021). Aging-related modular architectural reorganization of the metabolic brain network. *Brain Connect* Online ahead of print doi: 10.1089/brain.2021.0054

Ishibashi, K., Miura, Y., Toyohara, J., Ishii, K., and Ishiwata, K. (2017). Comparison of imaging using 11C-ITMM and 18F-FDG for the detection of cerebellar ataxia. *J. Neurol. Sci.* 375, 97–102. doi: 10.1016/j.jns.2017.01.036

Kakimoto, A., Ito, S., Okada, H., Nishizawa, S., Minoshima, S., Ouchi, Y., et al. (2016). Age-related sex-specific changes in brain metabolism and morphology. *J. Nucl. Med.* 57, 221–225. doi: 10.2967/jnumed.115.166439

Knopman, D. S., Jack, C. R., Wiste, H. J., Lundt, E. S., Weigand, S. D., Vemuri, P., et al. (2014). 18F-fluorodeoxyglucose positron emission tomography, aging, and apolipoprotein E genotype in cognitively normal persons. *Neurobiol. Aging* 35, 2096–2106. doi: 10.1016/j.neurobiolaging.2014.03.006

Malpetti, M., Ballarini, T., Presotto, L., Garibotto, V., Tettamanti, M., Perani, D., et al. (2017). Gender differences in healthy aging and Alzheimer's dementia: a 18F-FDG-PET study of brain and cognitive reserve. *Hum. Brain Mapp.* 38, 4212–4227. doi: 10.1002/hbm.23659

Mancho-Fora, N., Montalà-Flaquer, M., Farràs-Permanyer, L., Bartrés-Faz, D., Vaqué-Alcázar, L., Peró-Cebollero, M., et al. (2020). Resting-state functional dynamic connectivity and healthy aging: a sliding-window network analysis. *Psicothema* 32, 337–345. doi: 10.7334/psicothema2020.92

Moseley, M. (2002). Diffusion tensor imaging and AGING - A review. *NMR Biomed.* 15, 553–560.

Paldino, M. J., Zhang, W., Chu, Z. D., and Golriz, F. (2017). Metrics of brain network architecture capture the impact of disease in children with epilepsy. *NeuroImage Clin.* 13, 201–208. doi: 10.1016/j.nicl.2016.12.005

Ran, Q., Jamoulle, T., Schaeverbeke, J., Meersmans, K., Vandenberghe, R., and Dupont, P. (2020). Reproducibility of graph measures at the subject level using resting-state fMRI. *Brain Behav.* 10, 2336–2351. doi: 10.1002/brb3.1705

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003

Sala, A., and Perani, D. (2019). Brain molecular connectivity in neurodegenerative diseases: recent advances and new perspectives using positron emission tomography. *Front. Neurosci.* 13:617. doi: 10.3389/fnins.2019.00617

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., et al. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* 28, 3095–3114. doi: 10.1093/cercor/bhx179

Smith, C. D., Chebrolu, H., Wekstein, D. R., Schmitt, F. A., and Markesbery, W. R. (2007). Age and gender effects on human brain anatomy: a voxel-based morphometric study in healthy elderly. *Neurobiol. Aging* 28, 1075–1087. doi: 10.1016/j.neurobiolaging.2006.05.018

Sullivan, E. V., and Pfefferbaum, A. (2006). Diffusion tensor imaging and aging. *Neurosci. Biobehav. Rev.* 30, 749–761.

Sun, J., Tong, S., and Yang, G. Y. (2012). Reorganization of brain networks in aging and age-related diseases. *Aging Dis.* 3, 181–193.

Thomas, B. A., Erlandsson, K., Modat, M., Thurfjell, L., Vandenberghe, R., Ourselin, S., et al. (2011). The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. *Eur. J. Nucl. Med. Mol. Imaging* 38, 1104–1119. doi: 10.1007/s00259-011-1745-9

Van Aalst, J., Devrome, M., Van Weehaeghe, D., Rezaei, A., Radwan, A., Schramm, G., et al. (2021). Regional glucose metabolic decreases with ageing are associated with microstructural white matter changes: a simultaneous PET/MR study. *Eur. J. Nucl. Med. Mol. Imaging* Epub ahead of print. doi: 10.1007/s00259-021-05518-6

Varangis, E., Habeck, C. G., Razlighi, Q. R., and Stern, Y. (2019). The effect of aging on resting state connectivity of predefined networks in the brain. *Front. Aging Neurosci.* 11:234. doi: 10.3389/fnagi.2019.00234

Wang, Y., Ghumare, E., Vandenberghe, R., and Dupont, P. (2017). Comparison of different generalizations of clustering coefficient and local efficiency for weighted undirected graphs. *Neural Comput.* 29, 313–331. doi: 10.1162/NECO_a_00914

Wang, M., Jiang, J., Yan, Z., Alberts, I., Ge, J., Zhang, H., et al. (2020a). Individual brain metabolic connectome indicator based on Kullback-Leibler Divergence Similarity Estimation predicts progression from mild cognitive impairment to Alzheimer's dementia. *Eur. J. Nucl. Med. Mol. Imaging* 47, 2753–2764. doi: 10.1007/s00259-020-04814-x

Wang, M., Yan, Z., Xiao, S.-Y., Zuo, C., and Jiang, J. (2020b). A novel metabolic connectome method to predict progression to mild cognitive impairment. *Behav. Neurol.* 2020:2825037. doi: 10.1155/2020/2825037

Watabe, T., and Hatazawa, J. (2019). Evaluation of functional connectivity in the brain using positron emission tomography: a mini-review. *Front. Neurosci.* 13:775. doi: 10.3389/fnins.2019.00775

Yoshizawa, H., Gazes, Y., Stern, Y., Miyata, Y., and Uchiyama, S. (2014). Characterizing the normative profile of 18F-FDG PET brain imaging: sex difference, aging effect, and cognitive reserve. *Psychiatry Res.* 221, 78–85. doi: 10.1016/j.pscychresns.2013.10.009

Conflict of Interest: NM is Ph.D. mandate holder of the Fund for Scientific Research, Flanders, Belgium (FWO, 1S57420N). KVL has conducted this study as senior clinical research fellow for the Research Foundation Flanders (FWO).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 23 February 2022
doi: 10.3389/fnagi.2022.797803

Shunt Surgery Efficacy Is Correlated With Baseline Cerebrum Perfusion in Idiopathic Normal Pressure Hydrocephalus: A 3D Pulsed Arterial-Spin Labeling Study

Wenjun Huang[1†], Xuhao Fang[2†], Shihong Li[1], Renling Mao[2], Chuntao Ye[1], Wei Liu[1] and Guangwu Lin[1]*

[1] Department of Radiology, Huadong Hospital Affiliated to Fudan University, Shanghai, China, [2] Department of Neurosurgery, Huadong Hospital Affiliated to Fudan University, Shanghai, China

OPEN ACCESS

Edited by:
S. Abid Hussaini,
Columbia University Irving Medical Center, United States

Reviewed by:
John Robert Younce,
Washington University in St. Louis, United States
Derya Kaya,
Dokuz Eylül University, Turkey

***Correspondence:**
Guangwu Lin
lingw01000@163.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: 19 October 2021
Accepted: 24 January 2022
Published: 23 February 2022

Citation:
Huang W, Fang X, Li S, Mao R, Ye C, Liu W and Lin G (2022) Shunt Surgery Efficacy Is Correlated With Baseline Cerebrum Perfusion in Idiopathic Normal Pressure Hydrocephalus: A 3D Pulsed Arterial-Spin Labeling Study. Front. Aging Neurosci. 14:797803. doi: 10.3389/fnagi.2022.797803

This study investigated the relationship between preoperative cerebral blood flow (CBF) in patients with idiopathic normal pressure hydrocephalus (INPH) and preoperative clinical symptoms and changes of clinical symptoms after shunt surgery. A total of 32 patients with diagnosed INPH and 18 age-matched healthy controls (HCs) were involved in this study. All subjects underwent magnetic resonance imaging (MRI), including 3D pulsed arterial-spin labeling (PASL) for non-invasive perfusion imaging, and clinical symptom evaluation at baseline, and all patients with INPH were reexamined with clinical tests 1 month postoperatively. Patients with INPH had significantly lower whole-brain CBF than HCs, with the most significant differences in the high convexity, temporal lobe, precuneus, and thalamus. At baseline, there was a significant correlation between the CBF in the middle frontal gyrus, calcarine, inferior and middle temporal gyrus, thalamus, and posterior cingulate gyrus and poor gait manifestation. After shunting, improvements were negatively correlated with preoperative perfusion in the inferior parietal gyrus, inferior occipital gyrus, and middle temporal gyrus. Preoperative CBF in the middle frontal gyrus was positively correlated with the severity of preoperative cognitive impairment and negatively correlated with the change of postoperative MMSE score. There was a moderate positive correlation between anterior cingulate hypoperfusion and improved postoperative urination. Our study revealed that widely distributed and intercorrelated cortical and subcortical pathways are involved in the development of INPH symptoms, and preoperative CBF may be correlative to short-term shunt outcomes.

Keywords: idiopathic normal pressure hydrocephalus (INPH), pulsed arterial-spin labeling (PASL), neuroimaging, dementia, gait disorder, incontinence

INTRODUCTION

Normal-pressure hydrocephalus (NPH) is a hydrocephalus syndrome characterized by the clinical triad of gait disturbance, cognitive decline, and incontinence, with ventricular enlargement and a normal cerebrospinal fluid (CSF) pressure. Clinically, it can be classified as idiopathic normal pressure hydrocephalus (INPH) or secondary normal pressure hydrocephalus (SNPH) based

on whether the etiology is definite (Adams et al., 1965). At present, confirmation of INPH depends on the CSF tap test and shunt surgery (Nakajima et al., 2021), and clinical symptoms are treatable by permanent drainage (Relkin et al., 2005). However, a large number of patients with INPH are complicated with comorbidities, such as Alzheimer's disease (AD), Parkinson's-like diseases, extrapyramidal dyskinesia diseases (Allali et al., 2018), or similarities with normal aging (Agerskov et al., 2018), which may be the reason for a poor or even negatively impacting shunt response (Broggi et al., 2016; Bräutigam et al., 2019; Macki et al., 2020). With up to 40% of patients not responding, it remains difficult and essential to select patients who would benefit from surgery (Giordan et al., 2018). Therefore, by analyzing the relationship between baseline cerebral perfusion and changes in clinical symptoms after shunt, we attempted to infer the neural substrates underlying changes in clinical manifestations of INPH in order to screen the shunt surgery recipients and minimize unnecessary invasive procedures.

There is evidence that the pathological mechanism of INPH is complex and not only results from CSF circulation disorder (Lalou et al., 2018) but is also related to cerebrovascular self-regulation disorder and abnormal brain metabolism (Landau et al., 2014). Most studies have shown a significant decrease in regional cerebral blood flow (rCBF) in patients with INPH compared with healthy elderly individuals (Ziegelitz et al., 2014; Virhammar et al., 2017; Mattoli et al., 2020). Most studies have demonstrated that whole and regional cerebral blood flow (CBF) in patients with INPH were significantly lower than those in normal controls. Existing studies have not been consistent with the relationship between baseline CBF and change in clinical symptoms after shunt surgery. White matter CBF has been shown to gradually increase with greater distance from the ventricle, especially the lateral ventricle. In terms of gray matter, the frontal cortex and central gray matter were predominantly affected (Virhammar et al., 2014; Ziegelitz et al., 2016; Tuniz et al., 2017; Mattoli et al., 2020). Some studies suggest that hippocampal perfusion also decreased (Ziegelitz et al., 2014; Ziegelitz et al., 2015).

Arterial spin labeling (ASL) is a relatively new non-invasive perfusion imaging method that utilizes endogenous blood-based water molecules as tracers (without ionizing radiation) to visualize and quantify CBF (Jezzard et al., 2018). Based on the persistent stability of internal brain metabolic activity in the resting state, ASL can explore and monitor alterations in tissue perfusion in states of brain dysfunction (Soldozy et al., 2019).

The study aimed to determine the relationship between baseline CBF and preoperative clinical symptoms and the efficacy of shunt surgery. We hypothesized that baseline CBF may be associated with clinical changes after shunt surgery and help screen patients with INPH for shunt surgery clinically.

MATERIALS AND METHODS

Subjects

This study was approved by the Institutional Review Board of Huadong Hospital affiliated with Fudan University (approval number: 2017K027). The ethics committee waived the requirement of written informed consent for participation.

We retrospectively reviewed patients with INPH who were admitted to the inpatient unit at the neurosurgery department of Huadong Hospital affiliated with Fudan University to undergo shunt surgery from May 2019 to July 2021. The inclusion criterion for patients with diagnosed INPH according to expert consensus on the diagnosis and treatment of INPH in Experts consensus on diagnosis and treatment of normal pressure hydrocephalus in China (2016) were as follows: (1) age over 60 years; (2) the presence of at least one of the triad of symptoms (i.e., gait disturbance, dementia, or incontinence) with insidious progression for more than 6 months; (3) ventricular dilatation (Evans' index > 0.3); (4) CSF pressure < 200 mm H_2O; (5) the absence of other diseases that might account for such symptoms; and (6) underwent magnetic resonance examination, CSF tap test, and lumboperitoneal shunt surgery. The exclusion criteria for the patients with INPH were as follows: (1) cerebral infarction and dementia caused by clear causes and hospitalization for severe mental illness and (2) SNPH.

The inclusion criteria for elderly healthy controls (HCs) were as follows: (1) age over 60 years; (2) no gait disorder, cognitive impairment, or urination disorder, and normal Mini-Mental State Examination (MMSE) score; (3) conventional cerebral magnetic resonance imaging (MRI) showing no abnormalities; and (4) no active neurological, systemic, or psychiatric diseases.

Trained neurologists performed the clinical examinations. Besides a standard neurological examination, the tests included the INPH grading scale (INPHGS), MMSE, and for the patients with INPH, the timed up and go test (TUG-t) (Nakajima et al., 2021), before and 1 month after shunt surgery. For the INPHGS, motion disturbance, cognitive impairment, and incontinence are rated from 0 to 4. The higher the score was, the more severe the symptoms.

In total, we enrolled 32 patients with diagnosed INPH and 18 HCs into the study. The flowchart of enrollment of the diagnosed patients with INPH is shown in **Figure 1**. Demographic data and clinical characteristics are shown in **Table 1**.

Magnetic Resonance Protocol

All magnetic resonance (MR) data were acquired using a 3.0-T MRI scanner (MAGNETOM Prisma, Siemens Healthcare, Erlangen, Germany). The imaging protocol consisted of resting-state perfusion imaging assessment *via* a 3D pulsed arterial-spin labeling (PASL) sequence and anatomical assessment *via* T1-weighted magnetization prepared gradient-echo sequence (MPRAGE). A routine clinical T2-weighted structural MRI was also acquired for the neuroradiological assessment of the participants. For this study and based on preregistration, only the PASL and the MPRAGE sequences were included in the analyses.

The following parameters were used for the PASL acquisition: scan duration: 296 s; repetition time (TR): 4,600 ms; time to echo (TE): 16.18 ms; label time: 700 ms; post-labeling delay: 1,290 ms; inversion time (TI): 1,990 ms; field of view (FOV): 192 × 192; 3 mm × 3 mm × 3 mm; slice thickness: 3 mm with a 1.5 mm gap; 40 axial slices; and the number of excitations = 4.

The following parameters were used for the MPRAGE sequences: scan duration = 3 min 39 s; TR: 1,800 ms; TE: 2.37 ms;

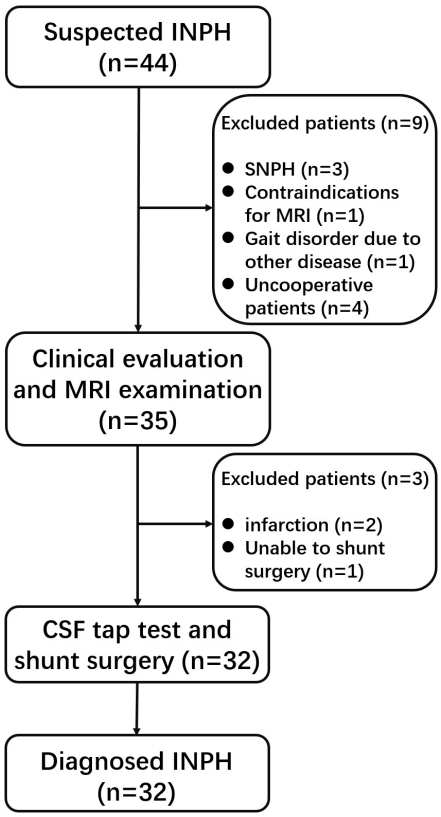


FIGURE 1 | Flowchart describing the inclusion process of patients.

FOV: 250 × 250; 0.87 mm × 0.85 mm × 0.85 mm; slice thickness: 0.85 mm with no gap; and 208 slices.

Magnetic Resonance Imaging Data Preprocessing

The PASL data were processed using the SPM12 software[1] and the ASL toolbox (ASLtbx[2]) (Wang et al., 2008). The center of each volume was first reset to the origin, and all rotations were set to zero. The first PASL image was set as the reference volume, and all other images were then motion-corrected relative to the reference. PASL images were realigned relative to the T1-weighted images for each subject. Smoothing of the realigned and coregistered PASL images was performed by applying an SPM Gaussian smoothing kernel of 6 mm × 6 mm × 6 mm full-width at half maximum (FWHM). A mask based on the mean of the smoothed PASL images was employed to exclude out-of-brain voxels. An ASL difference image was calculated using a single-compartment model (Buxton et al., 1998) after subtracting the label image from the control image. The four ASL difference images were averaged to calculate the CBF maps in combination with the proton-density-weighted reference images (Xu et al., 2010). Normalization, using T1 image unified segmentation with bounding box [–90, –126, –72; 93, 93, 111] and isotropic voxel size [3, 3, 3], could transform CBF maps to reduce the variability between individuals and allow meaningful group analyses. A quality check was performed visually to ensure the good quality of the preprocessing.

Statistical Analysis

Clinical symptom statistics are expressed as the median (quartile). We compared the data between the preoperative INPHs and HCs, and the measurements were made preoperatively and 1 month postoperatively using the Mann-Whitney *U*-test. The statistical analysis was conducted using the Statistical Package for Social Science version 24.0 (IBM SPSS).

Cerebral blood flow maps were statistically analyzed using second-level statistical procedures as implemented in SPM12 based on a generalized linear model (GLM). The CBF in the diagnosed INPH and HC groups was compared using a two-sample *t*-test. Multiple regression was used to analyze the correlation between preoperative CBF and preoperative and postoperative clinical changes by regressing out the *z*-transformed correlation coefficients.

Gender was included as covariates in the regression. The *p*-value $p < 0.05$ [false discovery rate (FDR) corrected, using the SPM12 software (see text footnote 1)] was considered significant. CBF clusters were visualized using the xjview[3] toolbox. Clusters

[1]http://www.fil.ion.ucl.ac.uk/spm

[2]https://cfn.upenn.edu/zewang/ASLtbx.php

[3]http://www.alivelearn.net/xjview

TABLE 1 | Demographic and clinical data in the diagnosed INPHs and HCs.

	INPH (*n* = 32)		Healthy control, HC (*n* = 18)	*p1*-value	*p2*-value
	Preoperative	1 month postoperative		Pre vs. HC	Pre vs. post
Age (average, range)	75.22 (66∼89)		72 (63∼82)	0.185	
Gender (male/female)	24/8		6/12	0.032*	
INPHGS					
Motion	3 (0)	2 (1)	0 (0.25)	< 0.001**	< 0.001**
Cognition	3 (1)	2 (0.75)	0 (1)	< 0.001**	< 0.001**
Urination	2.5 (1)	1 (1)	0 (0)	< 0.001**	< 0.001**
Total	8 (2)	5 (1)	0 (1.25)	< 0.001**	< 0.001**
MMSE	17.5 (8)	23 (7)	29 (2)	< 0.001**	0.037*
TUG-t	20.15 (10.09)	15.86 (3.85)	9.93 (1.58)	< 0.001**	0.001**

Values denote the median (quartile).
*Significant differences are marked with *p<0.05 and **p<0.01.*

with significant differences and significant correlations are displayed in pseudocolor on the calibrated standard brain map, and their Montreal Neurological Institute (MNI) coordinates and voxel sizes of peak intensity are listed in a table.

RESULTS

Demographic and Clinical Data

Table 1 shows the demographic and clinical variables of the patients with diagnosed INPH and HCs. No significant differences in age were observed between the patients with INPH and HCs ($p > 0.05$). As there were significant differences between the genders, we regressed it as a covariable in all statistical analyses.

The INPHGS, TUG, and MMSE scores markedly differed between the patients and HCs ($p < 0.001$). Furthermore, all patients with INPH improved in clinical symptoms to varying degrees after shunt surgery ($p < 0.05$).

Group Differences in Cerebral Blood Flow

Statistical analyses were observed regarding an automated anatomical atlas (AAL) template (Ashburner, 2007). The two-sample *t*-test revealed the significant differences between the patients with diagnosed INPH and HCs in the bilateral cerebrum. Generally, global and bilateral CBF was significantly lower in the INPHs than in the HCs, and the following brain areas were predominant: middle frontal gyrus (Frontal_Mid_L, R), thalamus, middle temporal gyrus (Temporal_Mid_R, L), precuneus (Precuneus_R, L), calcarine (Calcarine_L, R), inferior temporal gyrus (Temporal_Inf_L, R), corpus callosum, caudate, and middle cingulate gyrus (Cingulum_Mid_R, L) (**Figure 2**). The results were corrected by the FDR with a voxel-level $p < 0.05$.

Correlative Analysis

Regarding the diagnosed INPH group, we calculated Spearman correlation coefficients between the CBF values and preoperative clinical scale scores, including the INPHGS (motion, cognition,

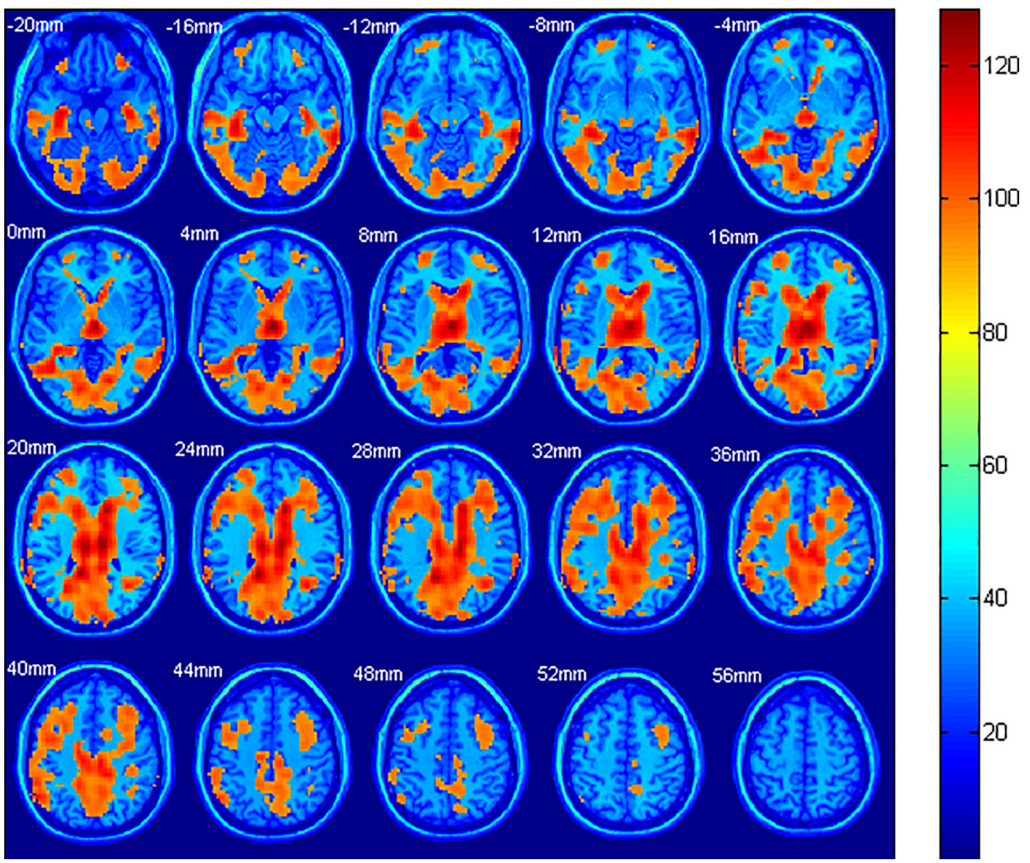


FIGURE 2 | Comparison of cerebral blood flow (CBF) among idiopathic normal pressure hydrocephalus (INPH) and healthy control (HC) groups. Significant region (false discovery rate corrected $p < 0.05$) illustrated in warm colors for increased values and in cool colors for decreased values. Significant differences were revealed in the following brain regions: bilateral middle frontal gyrus (Frontal_Mid_L, R), bilateral thalamus, bilateral middle temporal gyrus (Temporal_Mid_R, L), bilateral precuneus (Precuneus_R, L), bilateral calcarine (Calcarine_L, R), bilateral inferior temporal gyrus (Temporal_Inf_L, R), corpus callosum, bilateral caudate, and bilateral middle cingulate gyrus (Cingulum_Mid_R, L) ($p < 0.05$).

and urination), MMSE, and TUG-t. The brain regions related to the above clinical data are indicated in **Figures 3A–D** and **Table 2**. A positive correlation was identified between the following pairs: CBF values in the Frontal_Mid_L, R, Frontal_Sup_L, and Frontal_Inf_Oper_R and preoperative MMSE scores (**Figure 3A**). Negative correlations were identified between the following pairs: CBF values in the Frontal_Mid_R, Calcarine_R, L, Thalamus_R, L, and Cingulum_Post_L, R and preoperative TUG-t scores (**Figure 3B**), CBF values of the Occipital_Mid_L, Calcarine_L, Temporal_Inf_R, Temporal_Mid_R, and preoperative INPHGS-motion scores (**Figure 3C**), CBF values in the Frontal_Inf_Tri_R and preoperative INPHGS-cognition scores (**Figure 3D**) ($p < 0.001$).

Furthermore, we also calculated the correlation between the CBF values and postoperative clinical score changes, including the changes in INPHGS (motion, cognition, and urination), MMSE, and TUG-t scores. The brain regions related to the above clinical data are indicated in **Table 3**. Positive correlations were identified between the following pairs: CBF values in the Cingulum_Ant_R and postoperative INPHGS-urination score change (**Table 3**) and CBF values in the Temporal_Mid_R, L and postoperative TUG-t score change (**Table 3**). Negative correlations were identified between the following pairs: CBF values in the Frontal_Mid_R, L, Temporal_Mid_R, caudate, and thalamus and postoperative MMSE score change (**Table 3**) and CBF values in the Occipital_Inf_L, Parietal_Inf_L, Precuneus_R, and postoperative INPHGS-motion scores (**Table 3**) ($p < 0.001$).

DISCUSSION

In this study, PASL, a non-invasive perfusion imaging technique, was used to compare cerebral perfusion in patients with INPH with that in healthy older individuals. We found that the whole-brain CBF of patients with INPH was significantly reduced, which was consistent with the results of previous studies (Virhammar et al., 2014; Ziegelitz et al., 2016; Mattoli et al., 2020), suggesting a significant decrease in the cerebral perfusion rate. Hypoperfusion of the brain could further lead to pathophysiological changes in brain tissue, particularly oxygen metabolism (including oxygen utilization and extraction) (Göttler et al., 2019).

Overall, significantly reduced CBF in the INPH group was found in the high convexity gray matter. According to the *Guidelines for Management of Idiopathic Normal Pressure Hydrocephalus* (Third Edition), disproportionately enlarged subarachnoid space hydrocephalus (DESH) can be observed in most patients with INPH, which may be a reliable imaging feature of subarachnoid CSF absorption disorder (Nakajima et al., 2021). Reduced perfusion in the high convexity brain may result in local vascular compression due to its greater compression than in other regions of the brain. In addition, the CSF in the narrow sulcus is significantly reduced compared with other parts, so the metabolic efficiency of the local brain tissue is decreased.

There was reduced perfusion in the temporal lobe. One possible reason is the temporal lobe compression due to the temporal horn dilation, and another potential cause may be the AD comorbidity, which may also explain the precuneus hypoperfusion observed in this study (Ishii, 2020). Consistent with the findings of Ziegelitz's two studies on regional cerebral perfusion (Ziegelitz et al., 2014; Ziegelitz et al., 2015), CBF in the frontal lobe and periventricular white matter in patients with INPH was significantly lower than that in controls. Combined with previous studies (Ziegelitz et al., 2014; Virhammar et al., 2017), we speculated that the edema in the paraventricular white matter led to local compression of small vessels and metabolic disorder of vasoactive metabolites. This study found that thalamic

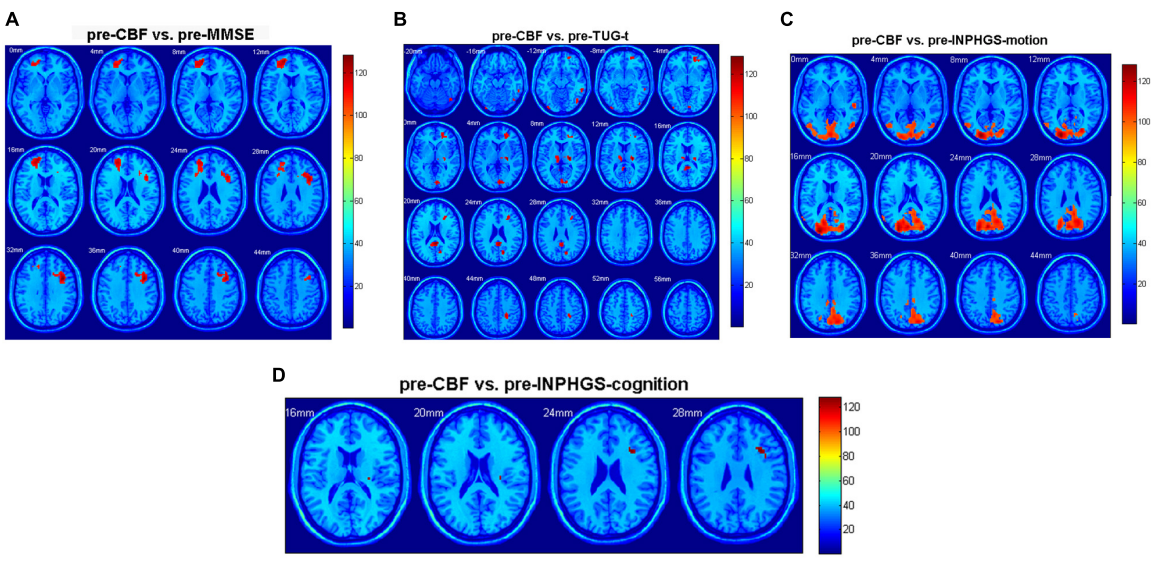


FIGURE 3 | Correlation of cerebral blood flow (CBF) between brain regions and clinical scale scores in patients with idiopathic normal pressure hydrocephalus (INPH): **(A)** Preoperative Mini-Mental State Examination (MMSE) score; **(B)** Preoperative timed up and go test (TUG-t) score; **(C)** Preoperative INPH grading scale (INPHGS) motion score; and **(D)** Preoperative INPHGS cognition score.

TABLE 2 | Correlation of cerebral blood flow (CBF) between brain regions and preoperative clinical scale score in idiopathic normal pressure hydrocephalus (INPH).

Preoperative clinical scale	No.	Brain region	Cluster size	Peak MNI coordinates (x, y, z)	Peak intensity
MMSE	1	Frontal_Mid_L Frontal_Sup_L	380	(−15, 54, 12)	4.71
	2	Frontal_Mid_R Frontal_Inf_Oper_R	300	(39, 9, 33)	5.18
TUG-t	1	Frontal_Mid_R	125	(18, 51, −6)	3.94
	2	Calcarine_R Calcarine_L	87	(3, −87, 3)	4.11
	3	Thalamus_L	45	(−9, −24, 9)	3.93
	4	Thalamus_R	61	(24, −18, 9)	4.83
	5	Cingulum_Post_L Cingulum_Post_R	120	(3, −45, 21)	4.69
INPHGS-motion	1	Occipital_Mid_L Calcarine_L	4145	(−21, −93, 12)	6.41
	2	Temporal_Inf_R Temporal_Mid_R	344	(54, −27, −3)	5.38
INPHGS-cognition	1	Frontal_Inf_Tri_R	21	(33, 18, 27)	3.79

TABLE 3 | Correlation of cerebral blood flow (CBF) between brain regions and postoperative clinical score changes in idiopathic normal pressure hydrocephalus (INPH).

Postoperative clinical score change	No.	Brain region	Cluster size	Peak MNI coordinates (x, y, z)	Peak intensity
MMSE	1	Frontal_Mid_R	49	(36, 21, 42)	3.83
	2	Caudate_R Thalamus_R	68	(6, −3, 6)	4.53
	3	Temporal_Mid_R	36	(45, −48, 15)	3.77
	4	Frontal_Mid_L	41	(−30, 9, 42)	4.35
TUG-t	1	Temporal_Mid_R	48	(54, −24, −6)	4.49
	2	Temporal_Mid_L	13	(−54, −39, 3)	3.61
INPHGS-motion	1	Occipital_Inf_L	70	(−21, −87, −3)	4.42
	2	Parietal_Inf_L	101	(−30, −60, 42)	4.02
	3	Precuneus_R	45	(3, −72, 45)	4.09
INPHGS-urination	1	Cingulum_Ant_R	16	(6, 33, 21)	3.73

CBF significantly decreased, which may be due to the decrease in N-acetyl-aspartate levels in the thalamus of patients with INPH due to impaired local metabolism (Miyamoto et al., 2007; Lundin et al., 2011), and downregulated circuits between cortical and subcortical structures, which may be related to the occurrence of clinical symptoms.

Many previous studies (Calamante, 2010; Ziegelitz et al., 2014, 2015) assumed that occipital cortex perfusion was less affected by diseases and used the occipital lobe as an internal reference to evaluate the perfusion in other brain regions. However, the results of this study showed that the occipital lobe perfusion in patients with INPH was significantly reduced compared with that in HCs, suggesting that there may be some deviation between the assumptions of previous studies and the real situation.

In this study, there was a significant correlation between the baseline CBF in the frontal lobe, temporal lobe, basal ganglia, thalamus, and cingulate gyrus and poor gait manifestation. Elderly individuals have more brain regions involved in motor control than younger ones. Previous studies found that the CBF in the frontal periventricular white matter was significantly correlated with gait (Jurcoane et al., 2014). The prefrontal cortex, which receives almost all information from the sensory system and is preferentially connected to motor information processing structures, plays a core role in cognitive control of motor performance, and therefore, elderly individuals are more dependent on activation of the bilateral frontal cortex during exercise. The frontal periventricular corticobasal ganglia-thalamocortical pathways are constituted by the fibers of the frontal lateral ventricle connected with the supplementary motor cortex, basal ganglia, and thalamus, which are involved in gait and body balance control (Virhammar et al., 2014). The thalamus and cingulate gyrus play important roles in this pathway, so reduced local perfusion may lead to downregulation of pathway function, which can manifest as motor dysfunction (Ziegelitz et al., 2015). The thalamus is the main structure regulating basal ganglia function and plays an important role in motor function. These structures influence each other and contribute to the occurrence and outcome of gait disorders in the course of diseases.

In the study, preoperative CBF in the frontal lobe and basal ganglia was positively correlated with the severity of preoperative cognitive impairment associated with INPH and negatively correlated with the change of postoperative MMSE

score. Decreased perfusion and metabolic disturbance lead to impaired ventricular white matter and then influence the function of the frontal-subcortical pathways and the frontal periventricular corticobasal ganglia-thalamocortical pathways, which further affects cognitive function. Previous studies have proven that this circuit is positively correlated with cognitive and psychological scores (Ziegelitz et al., 2016).

The improvement in motor function after shunting was negatively correlated with preoperative perfusion in the parietal lobe, occipital lobe, and temporal lobe. The occipital lobe is connected to the precuneus and frontal and temporal lobes through the cingulate tract and is involved in the regulation of spatial relations and visual attention of body movement (Tanglay et al., 2021). Some studies have shown that parietal occipital white matter damage is associated with gait disorder (de Laat et al., 2011). We hypothesized that it may be possible to predict the recovery of motor function after shunting by the perfusion damage observed in the parietal and occipitotemporal lobes before surgery.

The results showed that there was a moderate positive correlation between reduced anterior cingulate perfusion and improved postoperative urination. The higher brain's net effect on micturition is thought to be inhibitory. Moreover, the micturition reflex passes through the dorsolateral frontal cortex, anterior cingulate cortex, and hypothalamus (Tish and Geerling, 2020) as a part of the frontal-subcortical pathway and is closely related to the urinary control function. Preoperative perfusion in the cingulate gyrus being associated with recovery of bladder function is likely to predict the functional restoration of the descending cingulate pathway or cingulate cortex after shunt surgery.

Some studies have suggested that preoperative whole-brain CBF is related to clinical outcomes, and patients with lower CBF show clinical improvements after shunt surgery, indicating that preoperative CBF may contribute to predicting clinical outcomes after shunt surgery (Klinge et al., 2002). The slightly lower significance of our results compared with previous studies may be related to the small sample size and heterogeneity among the subjects, as well as differences in research methods and evaluation criteria of clinical symptoms. The results of previous studies have been discrepant, and whether baseline CBF can predict the outcome of shunt surgery has not been determined. Therefore, it is necessary to expand the sample size and conduct more in-depth research on this issue.

In some patients, motor function significantly improved in the short term after shunt surgery, but the duration was short, with the shortest maintenance of nearly a week. Specific analysis of individual patients revealed that a long course of the disease was a common feature. Combined with a CT perfusion (CTP) study (Ziegelitz et al., 2016), we speculated that massive CSF drainage could improve periventricular perfusion in the short term, but there might be permanent damage to white matter, which may interfere with the improvements in gray matter perfusion through the frontal periventricular corticobasal ganglia-thalamocortical pathways.

Positron emission tomography perfusion imaging, the only technique that can intrinsically quantify perfusion, is considered the gold standard in the cerebral perfusion evaluation with imaging (Mattoli et al., 2020). Previous studies have validated the CBF value of PASL against PET, proving that PASL has high repeatability in HCs and patients with AD (Xu et al., 2010). The advantage of the study lies in the quantitative analysis of CBF based on the whole-brain voxel level. On the one hand, this method directly analyzes the original data and does not involve *a priori* assumption of artificially defined ROI, so it is not subject to the subjective influence of researchers. On the other hand, the object of statistical analysis is each voxel in the CBF map, so the statistical result is not affected by volume. The spatial normalization process using the individuals' structural phase enables a voxel-based statistical comparison of brain images with different morphologies.

Compared with previous studies that take the occipital cortex as an internal reference to delineate regions of interest (ROIs), this study solved the limitations of manually drawing ROIs and avoided the influence of anatomical artifacts and volume on research results, resulting in higher accuracy. Therefore, voxel-based analysis has the advantages of automaticity, comprehensiveness, objectivity, and repeatability.

There were some limitations to this study that need to be considered. First, the post-labeling delay in the PASL technique used in this study was short (1,290 ms), which led to hypoperfusion artifacts in some subjects. Second, a single TI PASL was used in this study, which is less efficient and more dependent on model assumptions for the arterial transit time, resulting in slightly poorer accuracy. Considering that the subjects involved in this study could not undergo a long-time MRI scan due to cognitive impairment and old age, we chose the short post-labeling delay single TI PASL sequence as a trade-off. In the following study, we will continue to increase the sample size of the INPH group and strengthen follow-up, make a regression analysis of preoperative perfusion and clinical symptoms, and try to establish a prediction model of shunt efficacy to predict the efficacy through preoperative perfusion and help patients, their families, and healthcare professionals involved in treating INPH.

CONCLUSION

In this study, we measured brain perfusion in the patients with diagnosed INPH before and 1 month after shunt surgery to investigate the relationship between preoperative CBF and postoperative CBF changes and clinical symptoms. The current findings suggest that the HCs and the patients with diagnosed INPH exhibited CBF differences in the whole cerebrum, especially in the high convexity, temporal gyrus, and frontal white matter. The perfusion in different brain regions in the patients with INPH was correlated with clinical symptoms, and improvements in clinical symptoms after shunting were affected by the preoperative CBF. This study demonstrates that widely distributed and intercorrelated cortical and subcortical pathways are involved in the development of INPH symptoms. The pathogenesis of hypoperfusion and its specific effects on

disease development need to be further explored in combination with other imaging techniques and molecular studies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Huadong Hospital affiliated with Fudan University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

WH, XF, SL, RM, CY, WL, and GL made a substantial contribution to the concept and design, acquisition of data or analysis, and interpretation of data. WH, XF, and SL drafted the manuscript and revised it critically for relevant intellectual content. WH and XF performed the MR examination and follow-up of patients. All authors made a substantial contribution to the concept and design, acquisition of data or analysis, interpretation of data, and approved the final version of the manuscript.

FUNDING

This study was funded by the National Natural Science Foundation of China (81771816 to GL) and the Shanghai Municipal Population and Family Planning Commission (201740003 to RM).

REFERENCES

Adams, R. D., Fisher, C. M., Hakim, S., Ojemann, R. G., and Sweet, W. H. (1965). Symptomatic Occult Hydrocephalus With "Normal" Cerebrospinal-Fluid Pressure.A Treatable Syndrome. *New England J. Med.* 273, 117–126. doi: 10.1056/nejm196507152730301

Agerskov, S., Hellström, P., Andrén, K., Kollén, L., Wikkelsö, C., and Tullberg, M. (2018). The phenotype of idiopathic normal pressure hydrocephalus-a single center study of 429 patients. *J.Neurol. Sci.* 391, 54–60. doi: 10.1016/j.jns.2018.05.022

Allali, G., Laidet, M., Armand, S., and Assal, F. (2018). Brain comorbidities in normal pressure hydrocephalus. *Eur. J. Neurol.* 25, 542–548. doi: 10.1111/ene.13543

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage* 38, 95–113. doi: 10.1016/j.neuroimage.2007.07.007

Bräutigam, K., Vakis, A., and Tsitsipanis, C. (2019). Pathogenesis of idiopathic Normal Pressure Hydrocephalus: A review of knowledge. *J. Clin. Neurosci.* 61, 10–13. doi: 10.1016/j.jocn.2018.10.147

Broggi, M., Redaelli, V., Tringali, G., Restelli, F., Romito, L., Schiavolin, S., et al. (2016). Normal Pressure Hydrocephalus and Parkinsonism: Preliminary Data on Neurosurgical and Neurological Treatment. *World Neurosurg.* 90, 348–356. doi: 10.1016/j.wneu.2016.03.004

Buxton, R. B., Frank, L. R., Wong, E. C., Siewert, B., Warach, S., and Edelman, R. R. (1998). A general kinetic model for quantitative perfusion imaging with arterial spin labeling. *Magn. Reson. Med.* 40, 383–396. doi: 10.1002/mrm.1910400308

Calamante, F. (2010). Perfusion MRI using dynamic-susceptibility contrast MRI: quantification issues in patient studies. *Top. Magn. Reson. Imaging* 21, 75–85. doi: 10.1097/RMR.0b013e31821e53f5

de Laat, K. F., Tuladhar, A. M., van Norden, A. G., Norris, D. G., Zwiers, M. P., and de Leeuw, F. E. (2011). Loss of white matter integrity is associated with gait disorders in cerebral small vessel disease. *Brain* 134, 73–83. doi: 10.1093/brain/awq343

Experts consensus on diagnosis and treatment of normal pressure hydrocephalus in China. (2016). Experts consensus on diagnosis and treatment of normal pressure hydrocephalus in China. *Zhonghua Yi Xue Za Zhi* 96, 1635–1638.

Giordan, E., Palandri, G., Lanzino, G., Murad, M. H., and Elder, B. D. (2018). Outcomes and complications of different surgical treatments for idiopathic normal pressure hydrocephalus: a systematic review and meta-analysis. *J. Neurosurg.* 1, 1–13. doi: 10.3171/2018.5.Jns1875

Göttler, J., Preibisch, C., Riederer, I., Pasquini, L., Alexopoulos, P., Bohn, K. P., et al. (2019). Reduced blood oxygenation level dependent connectivity is related to hypoperfusion in Alzheimer's disease. *J. Cereb. Blood Flow Metab.* 39, 1314–1325. doi: 10.1177/0271678x18759182

Ishii, K. (2020). Diagnostic imaging of dementia with Lewy bodies, frontotemporal lobar degeneration, and normal pressure hydrocephalus. *Jpn. J. Radiol.* 38, 64–76. doi: 10.1007/s11604-019-00881-9

Jezzard, P., Chappell, M. A., and Okell, T. W. (2018). Arterial spin labeling for the measurement of cerebral perfusion and angiography. *J.Cereb. Blood Flow Metab.* 38, 603–626. doi: 10.1177/0271678x17743240

Jurcoane, A., Keil, F., Szelenyi, A., Pfeilschifter, W., Singer, O. C., and Hattingen, E. (2014). Directional diffusion of corticospinal tract supports therapy decisions in idiopathic normal-pressure hydrocephalus. *Neuroradiology* 56, 5–13. doi: 10.1007/s00234-013-1289-8

Klinge, P., Berding, G., Brinker, T., Schuhmann, M., Weckesser, E., Knapp, W. H., et al. (2002). The role of cerebral blood flow and cerebrovascular reserve capacity in the diagnosis of chronic hydrocephalus–a PET-study on 60 patients. *Acta Neurochir. Suppl.* 81, 39–41. doi: 10.1007/978-3-7091-6738-0_10

Lalou, A. D., Czosnyka, M., Donnelly, J., Pickard, J. D., Nabbanja, E., Keong, N. C., et al. (2018). Cerebral autoregulation, cerebrospinal fluid outflow resistance, and outcome following cerebrospinal fluid diversion in normal pressure hydrocephalus. *J. Neurosurg.* 130, 154–162. doi: 10.3171/2017.7.Jns17216

Landau, S. M., Thomas, B. A., Thurfjell, L., Schmidt, M., Margolin, R., Mintun, M., et al. (2014). Amyloid PET imaging in Alzheimer's disease: a comparison of three radiotracers. *Eur. J. Nucl. Med. Mol. Imaging* 41, 1398–1407. doi: 10.1007/s00259-014-2753-3

Lundin, F., Tisell, A., Dahlqvist Leinhard, O., Tullberg, M., Wikkelsö, C., Lundberg, P., et al. (2011). Reduced thalamic N-acetylaspartate in idiopathic normal pressure hydrocephalus: a controlled 1H-magnetic resonance spectroscopy study of frontal deep white matter and the thalamus using absolute quantification. *J. Neurol.Neurosurg.Psychiat.* 82, 772–778. doi: 10.1136/jnnp.2010.223529

Macki, M., Mahajan, A., Shatz, R., Air, E. L., Novikova, M., Fakih, M., et al. (2020). Prevalence of Alternative Diagnoses and Implications for Management in Idiopathic Normal Pressure Hydrocephalus Patients. *Neurosurgery* 87, 999–1007. doi: 10.1093/neuros/nyaa199

Mattoli, M. V., Treglia, G., Calcagni, M. L., Mangiola, A., Anile, C., and Trevisi, G. (2020). Usefulness of Brain Positron Emission Tomography with Different Tracers in the Evaluation of Patients with Idiopathic Normal Pressure Hydrocephalous. *Int. J. Mol. Sci.* 21:6523. doi: 10.3390/ijms21186523

Miyamoto, J., Imahori, Y., and Mineura, K. (2007). Cerebral oxygen metabolism in idiopathic-normal pressure hydrocephalus. *Neurol. Res.* 29, 830–834. doi: 10.1179/016164107X181851

Nakajima, M., Yamada, S., Miyajima, M., Ishii, K., Kuriyama, N., Kazui, H., et al. (2021). Guidelines for Management of Idiopathic Normal Pressure Hydrocephalus (Third Edition): Endorsed by the Japanese Society of Normal

Pressure Hydrocephalus. *Neurol. Med.Chir.* 61, 63–97. doi: 10.2176/nmc.st.2020-0292

Relkin, N., Marmarou, A., Klinge, P., Bergsneider, M., and Black, P. M. (2005). Diagnosing idiopathic normal-pressure hydrocephalus. *Neurosurgery* 57, S4–S16. doi: 10.1227/01.neu.0000168185.29659.c5

Soldozy, S., Galindo, J., Snyder, H., Ali, Y., Norat, P., Yağmurlu, K., et al. (2019). Clinical utility of arterial spin labeling imaging in disorders of the nervous system. *Neurosurg. Focus* 47:E5. doi: 10.3171/2019.9.Focus19567

Tanglay, O., Young, I. M., Dadario, N. B., Briggs, R. G., Fonseka, R. D., Dhanaraj, V., et al. (2021). Anatomy and white-matter connections of the precuneus. *Brain Imaging Behav.* [Epub Online ahead of print]. doi: 10.1007/s11682-021-00529-1

Tish, M. M., and Geerling, J. C. (2020). The Brain and the Bladder: Forebrain Control of Urinary (In)Continence. *Front. Physiol.* 11:658. doi: 10.3389/fphys.2020.00658

Tuniz, F., Vescovi, M. C., Bagatto, D., Drigo, D., De Colle, M. C., Maieron, M., et al. (2017). The role of perfusion and diffusion MRI in the assessment of patients affected by probable idiopathic normal pressure hydrocephalus. A cohort-prospective preliminary study. *Fluids Barriers CNS* 14:24. doi: 10.1186/s12987-017-0072-3

Virhammar, J., Laurell, K., Ahlgren, A., Cesarini, K. G., and Larsson, E. M. (2014). Idiopathic normal pressure hydrocephalus: cerebral perfusion measured with pCASL before and repeatedly after CSF removal. *J. Cereb. Blood Flow Metab.* 34, 1771–1778. doi: 10.1038/jcbfm.2014.138

Virhammar, J., Laurell, K., Ahlgren, A., and Larsson, E. M. (2017). Arterial Spin-Labeling Perfusion MR Imaging Demonstrates Regional CBF Decrease in Idiopathic Normal Pressure Hydrocephalus. *AJNR Am. J. Neuroradiol.* 38, 2081–2088. doi: 10.3174/ajnr.A5347

Wang, Z., Aguirre, G. K., Rao, H., Wang, J., Fernández-Seara, M. A., Childress, A. R., et al. (2008). Empirical optimization of ASL data analysis using an ASL data processing toolbox: ASLtbx. *Mag. Reson. Imaging* 26, 261–269. doi: 10.1016/j.mri.2007.07.003

Xu, G., Rowley, H. A., Wu, G., Alsop, D. C., Shankaranarayanan, A., Dowling, M., et al. (2010). Reliability and precision of pseudo-continuous arterial spin labeling perfusion MRI on 3.0 T and comparison with 15O-water PET in elderly subjects at risk for Alzheimer's disease. *NMR Biomed.* 23, 286–293. doi: 10.1002/nbm.1462

Ziegelitz, D., Arvidsson, J., Hellström, P., Tullberg, M., Wikkelsø, C., and Starck, G. (2015). In Patients With Idiopathic Normal Pressure Hydrocephalus Postoperative Cerebral Perfusion Changes Measured by Dynamic Susceptibility Contrast Magnetic Resonance Imaging Correlate With Clinical Improvement. *J. Comp. Assist. Tomogr.* 39, 531–540. doi: 10.1097/rct.0000000000000254

Ziegelitz, D., Arvidsson, J., Hellström, P., Tullberg, M., Wikkelsø, C., and Starck, G. (2016).). Pre-and postoperative cerebral blood flow changes in patients with idiopathic normal pressure hydrocephalus measured by computed tomography (CT)-perfusion. *J. Cereb. Blood Flow Metab.* 36, 1755–1766. doi: 10.1177/0271678x15608521

Ziegelitz, D., Starck, G., Kristiansen, D., Jakobsson, M., Hultenmo, M., Mikkelsen, I. K., et al. (2014). Cerebral perfusion measured by dynamic susceptibility contrast MRI is reduced in patients with idiopathic normal pressure hydrocephalus. *J. Magn. Reson. Imaging* 39, 1533–1542. doi: 10.1002/jmri.24292

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

CASE REPORT
published: 27 April 2022
doi: 10.3389/fnagi.2022.804074

Case Report: Advantages of High-Resolution MRI in Evaluating the Efficacy of Drug Therapy for Intracranial Atherosclerotic Plaques

Tao Zheng[1,2], *Lei Liu*[2], *Li Li*[3], *Yang Gao*[2], *Runcai Guo*[4], *Zhi Zhou*[2], *Zunjing Liu*[2]* and *Kunpeng Liu*[5]*

[1] Graduate School, Beijing University of Chinese Medicine, Beijing, China, [2] Department of Neurology, China-Japan Friendship Hospital, Beijing, China, [3] Department of Anesthesiology, Beijing Friendship Hospital, Capital Medical University, Beijing, China, [4] Department of Radiology, China-Japan Friendship Hospital, Beijing, China, [5] Department of Anesthesiology, Peking University International Hospital, Beijing, China

Intracranial atherosclerotic stenosis is one of the main causes of ischemic stroke and transient ischemic attack. High-resolution magnetic resonance imaging allows us to directly observe the intracranial artery wall, accurately assess the condition of the vascular wall, and quantitatively analyze the vascular wall and intracranial atherosclerotic plaque load. We report a case of acute cerebral infarction with left middle cerebral artery stenosis. During the first 3 weeks, the patient was treated with aspirin 100 mg and clopidogrel 75 mg daily. Afterwards, the patient continued to be given aspirin, and cilostazol 100 mg twice daily was given instead of clopidogrel. After 24 months of follow-up, we observed a significant reversal of intracranial atherosclerotic plaque using high-resolution MRI (HR-MRI) and discussed the advantages of HR-MRI in evaluating drug therapy for intracranial atherosclerotic plaque.

Keywords: high-resolution MRI, advantages, evaluating, intracranial atherosclerotic plaques, reversal

OPEN ACCESS

Edited by:
Yang Jiang,
University of Kentucky, United States

Reviewed by:
Zhen-Ni Guo,
First Affiliated Hospital of Jilin University, China
Hyo Sung Kwak,
Chonbuk National University Hospital, South Korea

***Correspondence:**
Zunjing Liu
liuzunjing@163.com
Kunpeng Liu
liukunpeng@pkuih.edu.cn

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *28 October 2021*
Accepted: *28 March 2022*
Published: *27 April 2022*

Citation:
Zheng T, Liu L, Li L, Gao Y, Guo R, Zhou Z, Liu Z and Liu K (2022) Case Report: Advantages of High-Resolution MRI in Evaluating the Efficacy of Drug Therapy for Intracranial Atherosclerotic Plaques. Front. Aging Neurosci. 14:804074. doi: 10.3389/fnagi.2022.804074

INTRODUCTION

Intracranial atherosclerotic stenosis (ICAS) is one of the main causes of ischemic stroke and transient ischemic attack (TIA). Stroke is also a major cause of death and disability worldwide (Rahman et al., 2021). Active intervention in plaque progression and prevention of stroke are the main objectives of current drug therapy. Most existing studies evaluate the efficacy of drugs in the treatment of intracranial atherosclerotic plaques based only on changes in lumen, while few studies evaluate the plaques' diameter of the stenosis, vascular area, lumen area, wall area, stenosis rate, remodeling index, wall area index, and normalized wall index using HR-MRI (Chen et al., 2018; Chung et al., 2020; Shi et al., 2021). We report a case of acute cerebral infarction with left middle cerebral artery stenosis. During the first 3 weeks, the patient was treated with aspirin 100 mg and clopidogrel 75 mg daily. Afterwards, the patient continued to be given aspirin, and cilostazol 100 mg twice daily was given instead of clopidogrel. After 24 months of follow-up, we observed a significant reversal of intracranial atherosclerotic plaque using high-resolution MRI (HR-MRI) and discussed the advantages of HR-MRI in evaluating drug therapy for intracranial atherosclerotic plaque.

CASE DESCRIPTION

A 46-year-old Chinese female presented with a sudden-onset right-sided weakness and numbness 1 day ago. She denied diplopia, visual field change, disturbance of consciousness, headache, and

dizziness. She also denied any history of hypertension, hyperlipidemia, and others. There are no obvious abnormalities in the patient's personal and family history. In the emergency room, the patient's blood pressure was 163/85 mmHg. On neurological examination, the muscle strength was IV, the sensation of pinprick was impaired, and Babinski sign was positive in the right limb. The routine laboratory examinations after admission showed that low density lipoprotein cholesterol was 3.47 mmol/L, triglycerides was 1.79 mmol/L, and the rest of the routine laboratory examinations showed no obvious abnormalities. Head MRI showed multiple small areas of acute infarction on the left basal ganglia, thalamus, and temporal lobe (**Figures 1A–F**). During the first 3 weeks, the patient was treated with aspirin 100 mg daily and clopidogrel 75 mg daily for antiplatelet therapy. Afterwards, the patient continued to be given aspirin, and cilostazol 100 mg twice daily was given instead of clopidogrel. The reason for using cilostazol was that we referred to the CSPS.com trial (see Section Discussion) (Toyoda et al., 2019). During the follow-up, rosuvastatin 10 mg daily and amlodipine 10 mg daily was continued. During the follow-up, the lipid control objectives were as follows: LDL-C $\leq$ 1.8 mmol/L, systolic blood pressure$\leq$140 mmHg, without complications or neurological deficits. HR-MRI examinations at hospitalization showed moderate to severe stenosis of the M1 section of the left middle cerebral artery (MCA; **Figures 2A–C**). The reexamination result of HR-MRI at 24 months was better, during which the lumen was slightly narrowed and the wall was slightly thickened in the same place (**Figures 2D–F**).

METHODS

The high-resolution MRI used in our study was a 3.0 T MRI scanner (Ingenia; Philips Healthcare, Best, The Netherlands) with a 15-channel phased-array coil. Our imaging results were diagnosed by two radiologists with more than 5 years of experience in HR-MRI diagnosis. Three-dimensional (3D) volume isotropic turbo spin-echo acquisition (VISTA) images were obtained by axial plane scanning of the major intracranial arteries with the following parameters: repetition time/echo time, 1,300/36 ms; field of view, 140 × 200 × 105 mm^3; matrix, 280 × 332 × 210; number of excitations, 2. Acquisition voxel volume was 0.5× 0.6 × 0.5 mm^3, and reconstruction voxel volume was 0.5 × 0.5 × 0.5 mm^3. Axial plane images were automatically constructed with a slice thickness of 0.5 mm. The 3D-VISTA scan time was approximately 5 min. We selected time of flight (TOF) magnetic resonance angiography (MRA) images and 3D-VISTA images. MRA scan was mainly used to determine the location and degree of MCA stenosis and to angle the reconstruction

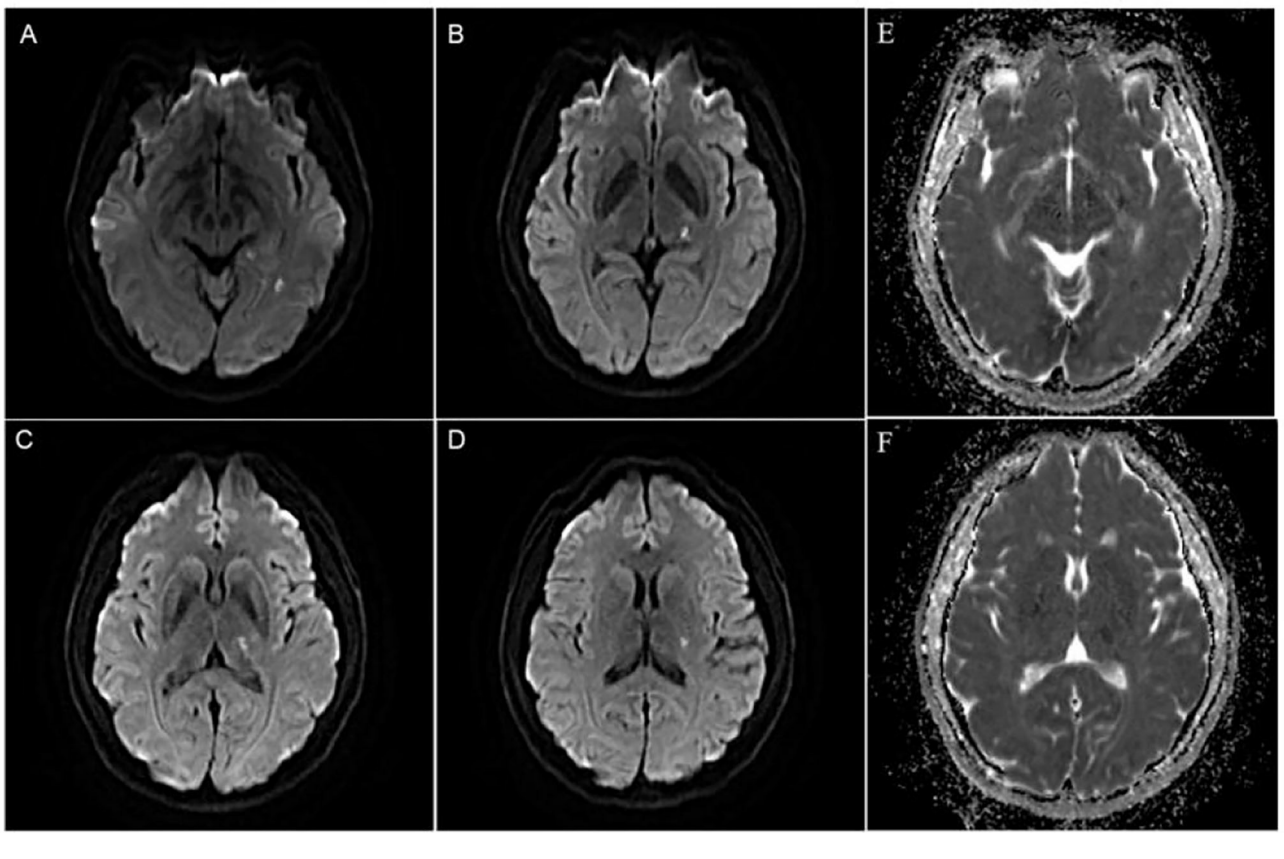


FIGURE 1 | (A–D) Diffusion weighted imaging(DWI) showed multiple and small areas of acute infarction on the left basal ganglia, thalamus, and temporal lobe. **(E,F)** ADC image showed multiple and small areas of low signal areas on the left basal ganglia, thalamus, and temporal lobe.

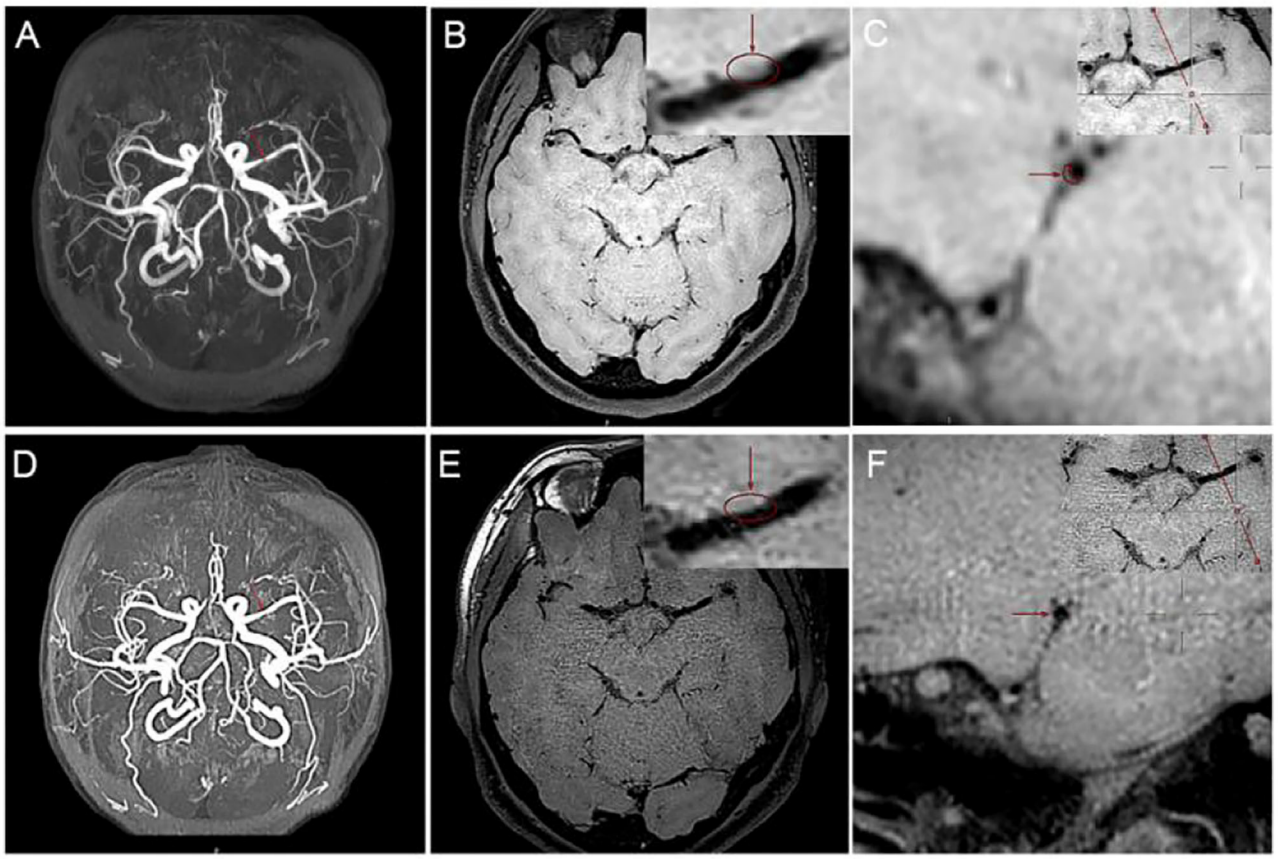


FIGURE 2 | High-resolution MRI (HR-MRI) examination at hospitalization. **(A)** Time of flight (TOF) magnetic resonance angiography (MRA) showed severe M1 stenosis in the left middle cerebral artery (MCA). **(B)** T1-weighted volume isotropic turbo spin-echo acquisition (T1W-VISTA) showed the location of the plaque and thickens the eccentric canal wall. T1W-VISTA also showed no significant enhancement on the plaque or surface, similar to the surrounding tissue signals, indicating no thrombosis on the surface of the plaque. **(C)** 3D-VISTA reconstruction was carried out at the maximal-lumen-narrowing site, and the rough shape of the patch was hand-drawn. HR-MRI examination after 24 months. **(D)** TOF MRA showed that the degree of lumen stenosis was significantly reduced. **(E)** T1W-VISTA showed that there were still some plaques on the tube wall, and the lumen was slightly narrow. **(F)** 3D-VISTA reconstruction showed the diameter of the tube was significantly larger.

plane of 3D-VISTA images to ensure that all cross-sections were perpendicular to the long axis of the MCA. The cross section of the MCA was divided into four quadrants: upper, lower, ventral, and dorsal quadrant. Moreover, changing the color map of the image from the traditional gray to sky blue will make the outline and thickness of the blood vessel wall appear more clearly. As shown in **Figure 3**, after we changed the color mapping of digital picture archiving and communication (PACS) workstation from gray scale to sky blue, the contour of blood vessel wall in **Figure 3B** was clearer than in **Figure 3A**.

We semi-automatically obtained the vessel diameter, wall thickness, vessel area (VA), and lumen area (LA) of the maximal-lumen-narrowing (MLN) site and the reference site by using the measurement tool on PACS. The calculation formula of the wall area (WA) was WA=VA-LA. The vessels in the non-stenosis segment of the proximal MCA were used as reference points, and the reference wall area and lumen area were measured. The degree of stenosis was computed as follows: degree of stenosis = $(1 - LA_{MLN} / LA_{reference}) \times 100\%$. The remodeling index (RI) was defined as $VA_{MLN} / VA_{reference}$. The wall area index (WAI) was defined as $WA_{MLN} / WA_{reference}$. The normalized wall index (NWI) was defined as $WA_{MLN} / (WA_{MLN}+LA_{MLN}) \times 100\%$. RI ≤ 0.95 was defined as negative remodeling (NR). Lastly, $0.95 < RI < 1.05$ was defined as intermediate remodeling (IR), and RI ≥ 1.05 was defined as positive remodeling (PR).

RESULTS

After 24 months of follow-up, we found that the diameter of the stenosis increased from 1.08 to 1.90 mm, the LA increased from 1.13 to 2.56 mm^2, the WA decreased from 4.35 to 2.92 mm^2, the stenosis rate on HR-MRI decreased from 49.53 to 11.21%, the RI increased from 39.65 to 89.82%, the WAI decreased from 1.14 to 0.76, and the NWI decreased from 79.38 to 53.28%, as shown in **Table 1**.

DISCUSSION

How to more accurately and more safely evaluate the efficacy of drug therapy for intracranial atherosclerotic plaque is

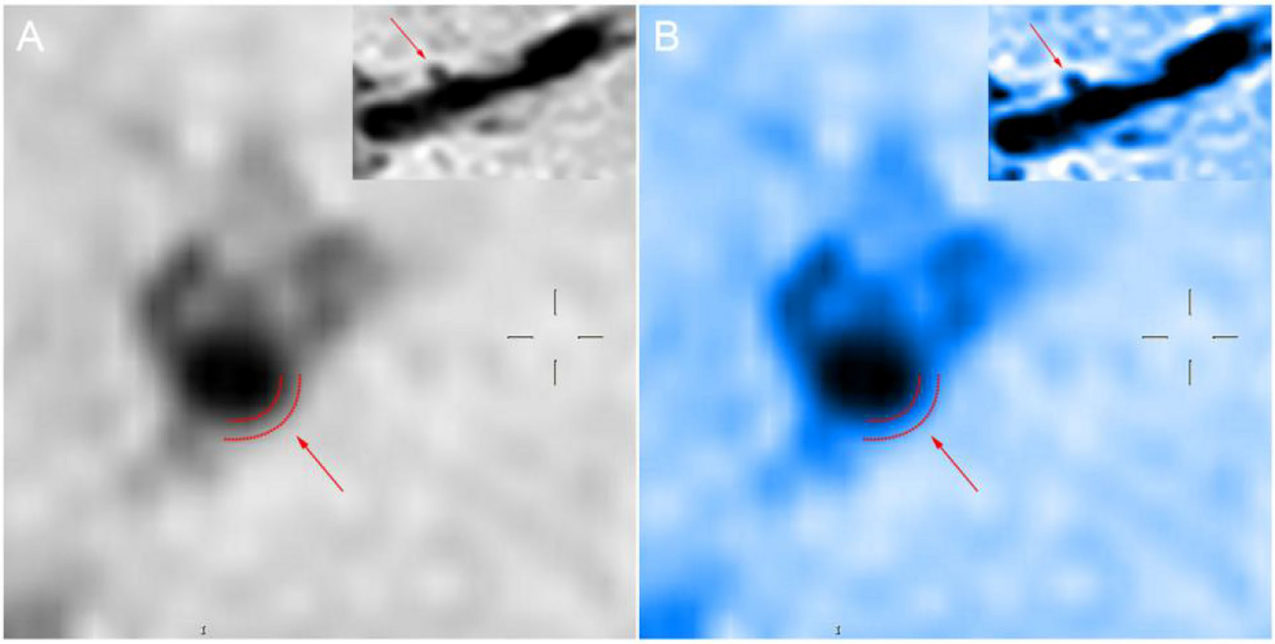


FIGURE 3 | Panels **(A,B)** show the sagittal position of the perforating artery as shown by the arrow at the upper right. The boundary delineated by the dashed line indicated by the arrow in the larger image is the rough outline of the vessel wall.

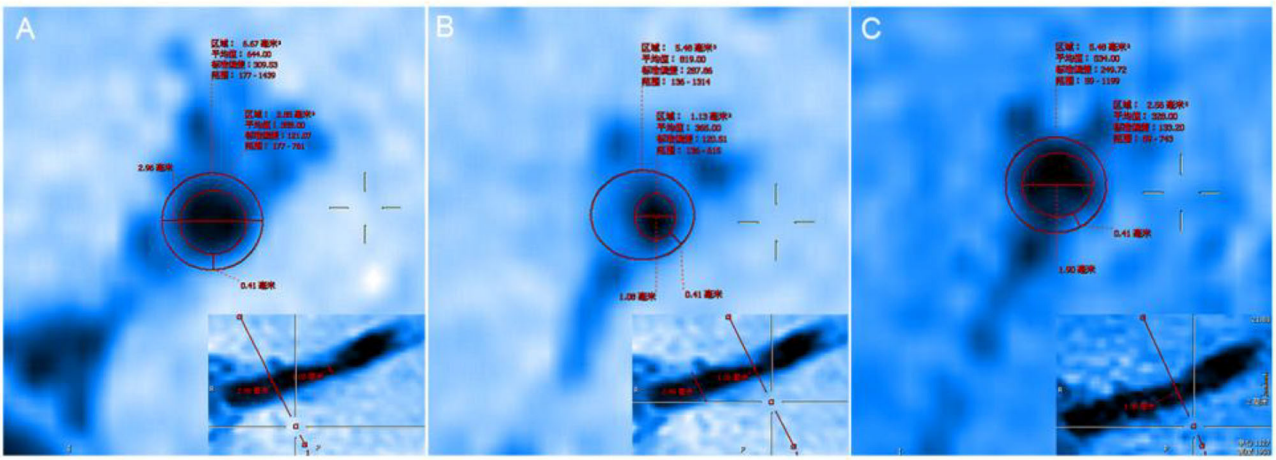


FIGURE 4 | (A) The proximal vessels serve as the reference site. **(B)** The maximal-lumen-narrowing site before follow-up. **(C)** The maximal-lumen-narrowing site after follow-up.

one of the problems that need to be solved in clinical practice. CT angiography (CTA), MRA, and digital subtraction angiography (DSA) help us assess intracranial atherosclerotic stenosis by primarily providing residual lumen diameter. The methods mentioned above do not provide us with information about vascular walls, atherosclerotic plaques, etc. HR-MRI can directly observe the vascular wall and lumen and perform quantitative and qualitative analysis of plaque, thus providing more information of clinical value. MRA may exaggerate the extent of lumen narrowing. For example, the arrow in **Figure 2A** shows severe stenosis in the M1 segment of the left MCA on MRA, whereas **Figure 2B** shows approximately moderate stenosis on HR-MRI. Through quantitative analysis, we calculated that the stenosis of the left MCA was approximately 49.53%, which was approximately moderate and not far from the stenosis we estimated on T1W-VISTA compared to the severe stenosis observed on MRA. This also confirms that MRA tend to overestimate the stenosis of blood vessels. Previous studies have

TABLE 1 | Changes of various indexes before and after follow-up.

	D (mm)	VA (mm^2)	LA (mm^2)	WA (mm^2)	R (%)	RI (%)	WAI	NWI (%)
HR-MRI at hospitalization	1.08	5.48	1.13	4.35	49.53	39.65	1.14	79.38
HR-MRI after 24 months	1.90	5.48	2.56	2.92	11.21	89.82	0.76	53.28

HR-MRI, High-resolution MRI; D, The diameter of the maximal-lumen-narrowing site; VA, Vascular area; LA, Lumen area; WA, Wall area; R, Stenosis rate; RI, Remodeling index; WAI, Wall area index; NWI, Normalized wall index.

shown that CTA may also exaggerate lumen stenosis (Liu et al., 2013).

In evaluating the effects of drug therapy on intracranial atherosclerotic plaque, the changes of plaque and vascular wall should be paid attention to in addition to lumen. However, MRA, CTA, and DSA cannot provide us with information about plaques and vascular walls. In addition to stenosis, HR-MRI can also measure wall thickness, plaque volume, LA, WA, RI, WAI, NWI, etc. WAI and NWI are commonly used indexes to analyze plaque load, and their calculation methods are shown in the method section of this paper. Large culprit plaque load was independently associated with recurrent acute stroke (Sun et al., 2021). Plaque load can also be used to assess plaque size and drug efficacy. We calculated the WAI and NWI to compare plaque load before and after follow-up, as shown in **Figure 4**. The WAI decreased from the initial value of 1.14 to 0.74, and the NWI was approximately 79.38% before and 53.28% after the follow-up, significantly decreased by 26.1%. All the above findings indicate that the plaque load was significantly reduced under the intervention of drugs. Although both methods can describe plaque load, the latter is increasingly accepted by the majority of scholars and is regarded as one of the best indicators for evaluating plaque load at present. The main reason is that WAI is calculated according to the ratio of the WA of the MLN site to the WA of the proximal non-stenosis site, which is relatively subjective. However, NWI is based on the ratio of the wall area of the MLN site, which is relatively objective and has certain reference value. The RI is calculated by the wall area, which can be used to judge the stability of plaque. It is mainly divided into positive remodeling and negative remodeling. Positive remodeling is often associated with plaque instability and is prone to cause acute ischemic stroke (Zhang et al., 2017). For stroke patients with positive remodeling, luminal imaging at this time often shows normal or mild stenosis, and the pathological mechanism of stroke is difficult to explain. HR-MRI can help us study the mechanism of infarction in such patients with positive remodeling (Sun et al., 2018).

The duration of treatment and the combination of drugs may be correlated with whether the vascular remodeling index changes significantly. In Chung and Shi's study, significant changes in the vascular RI were not found (Chung et al., 2020; Shi et al., 2021). Particularly, Chung found no significant change in the RI after intensive statin treatment (1.09 ± 0.35 vs. 1.03 ± 0.30; $p = 0.195$) (Chung et al., 2020). In our study, the vascular RI before follow-up was 0.3965, which was much less than 0.95. After follow-up, the new RI was 0.8982, which was very close to the range of normal vascular RI. Prolonged statin therapy may make the plaque more stable and may also bring the remodeling index back into the normal range. The reason may be that the duration of intensive statin therapy (6 months) observed in Chung's study was relatively short. Long-term and appropriate antiplatelet therapy is an indispensable link to reduce clinical events. This was a patient with symptomatic intracranial artery stenosis who is at high risk of recurrent ischemic stroke. However, the patient's DWI showed a small ischemic focus, and the patient was at a low risk of bleeding after taking antiplatelet drugs. The reason we chose cilostazol and aspirin long-term combination therapy is that we referred to the "CSPS.com" trial in Japan, which showed that for high-risk stroke patients, long-term combination therapy of cilostazol with aspirin or clopidogrel had a lower risk of ischemic stroke recurrence (Toyoda et al., 2019). There was no significant increase in the risk of serious or life-threatening bleeding. Other studies confirm this (Lin et al., 2021). Previous studies have found that cilostazol provides benefits after 60–90 days of treatment (de Havenon et al., 2021). In addition to inhibiting platelet aggregation, cilotazol can also promote blood vessel remodeling and lowering of blood lipid levels, among others. Some studies found that cilostazol can stabilize or reverse plaque effects. For example, Lee et al. found that after long-term treatment with cilostazol, carotid plaque volume and intima thickness were significantly reduced, leading to vascular remodeling (Lee et al., 2020). In our study, the reversal of plaque in patients may be related to the use of cilostazol.

CONCLUSION

High-resolution MRI is one of the more intuitive and non-invasive auxiliary examinations. Through quantitative analysis of vascular wall and plaque, HR-MRI can help us compare the effects of drugs on delaying or reversing intracranial atherosclerotic plaque from multiple perspectives.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of clinical trials of drugs or devices in China-Japan Friendship Hospital. The

patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TZ contributed to paper conception, data collection, analysis and interpretation, literature review, and paper drafting. LLiu, LLi, YG, RG, and ZZ contributed to data collection, analysis and interpretation, literature review, and paper drafting. ZL and KL contributed to paper conception, the supervision of data analysis and interpretation, and critically reviewed the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (No. 52073310) and Elite Medical Professionals project of China-Japan Friendship Hospital (NO. ZRJY2021-BJ03).

REFERENCES

Chen, X., Wang, S., Lin, L., Li, Y., and Zhang, H. (2018). Drug effect of atorvastatin on middle cerebral atherosclerotic stenosis and high resolution NMR diagnosis. *Pak. J. Pharm. Sci.* 31, 1169–1173.

Chung, J. W., Cha, J., Lee, M. J., Yu, I. W., Park, M. S., Seo, W. K., et al. (2020). Intensive statin treatment in acute ischaemic stroke patients with intracranial atherosclerosis: a high-resolution magnetic resonance imaging study (STAMINA-MRI study). *J. Neurol. Neurosurg. Psychiatry* 91, 204–211. doi: 10.1136/jnnp-2019-320893

de Havenon, A., Sheth, K. N., Madsen, T. E., Johnston, K. C., Turan, T. N., Toyoda, K., et al. (2021). Cilostazol for secondary stroke prevention: history, evidence, limitations, and possibilities. *Stroke* 52, e635–e645. doi: 10.1161/STROKEAHA.121.035002

Lee, D. H., Chun, E. J., Moon, J. H., Yun, H. M., and Lim, S. (2020). Effect of cilostazol on carotid plaque volume measured by three-dimensional ultrasonography in patients with type 2 diabetes: the FANCY study. *Diabetes Obes. Metab.* 22, 2257–2266. doi: 10.1111/dom.14147

Lin, M. P., Meschia, J. F., Gopal, N., Barrett, K. M., Ross, O. A., Ertekin-Taner, N., et al. (2021). Cilostazol versus aspirin for secondary stroke prevention: systematic review and meta-analysis. *J. Stroke Cerebrovasc. Dis.* 30, 105581. doi: 10.1016/j.jstrokecerebrovasdis.2020.105581

Liu, Q., Huang, J., Degnan, A. J., Chen, S. Y., Gillard, J. H., Teng, Z. Z., et al. (2013). Comparison of high-resolution MRI with CT angiography and digital subtraction angiography for the evaluation of middle cerebral artery atherosclerotic steno-occlusive disease. *Int. J. Cardiovasc. Imaging* 29, 1491–1498. doi: 10.1007/s10554-013-0237-3

Rahman, A. A., Amruta, N., Pinteaux, E., and Bix, G. J. (2021). Neurogenesis after stroke: a therapeutic perspective. *Transl. Stroke Res.* 12, 1–14. doi: 10.1007/s12975-020-00841-w

Shi, Z., Li, J., Zhao, M., Zhang, X., Degnan, A. J., Mossa-Basha, M., et al. (2021). Progression of plaque burden of intracranial atherosclerotic plaque predicts recurrent stroke/transient ischemic attack: a pilot follow-up study using higher-resolution MRI. *J. Magn. Reson. Imaging* 54, 560–570. doi: 10.1002/jmri.27561

Sun, B., Wang, L., Li, X., Zhang, J., Zhang, J., Liu, X., et al. (2021). Intracranial atherosclerotic plaque characteristics and burden associated with recurrent acute stroke: a 3D quantitative vessel wall MRI study. *Front. Aging Neurosci.* 13, 706544. doi: 10.3389/fnagi.2021.706544

Sun, L. L., Li, Z. H., Tang, W. X., Liu, L., Chang, F. Y., Zhang, X. B., et al. (2018). High resolution magnetic resonance imaging in pathogenesis diagnosis of single lenticulostriate infarction with nonstenotic middle cerebral artery, a retrospective study. *BMC Neurol.* 18, 51. doi: 10.1186/s12883-018-1054-z

Toyoda, K., Uchiyama, S., Yamaguchi, T., Easton, J. D., Kimura, K., Hoshino, H., et al. (2019). Dual antiplatelet therapy using cilostazol for secondary prevention in patients with high-risk ischaemic stroke in Japan: a multicentre, open-label, randomised controlled trial. *Lancet Neurol.* 18, 539–548. doi: 10.1016/S1474-4422(19)30148-6

Zhang, D. F., Chen, Y. C., Chen, H., Zhang, W. D., Sun, J., Mao, C. N., et al. (2017). A high-resolution MRI study of relationship between remodeling patterns and ischemic stroke in patients with atherosclerotic middle cerebral artery stenosis. *Front. Aging Neurosci.* 9, 140. doi: 10.3389/fnagi.2017.00140

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 30 May 2022
doi: 10.3389/fnagi.2022.904194

Dual-Task Performance in Older Adults With and Without Idiopathic Normal Pressure Hydrocephalus

Otto Lilja-Lund[1], Lars Nyberg[2,3,4,5], Martin Maripuu[6] and Katarina Laurell[7]*

[1] Department of Clinical Sciences, Neuroscience, Umeå University, Umeå, Sweden, [2] Department of Radiation Sciences, Radiology, Umeå University, Umeå, Sweden, [3] Department of Integrative Medical Biology, Umeå University, Umeå, Sweden, [4] Umeå Center for Functional Brain Imaging, Umeå University, Umeå, Sweden, [5] Center for Lifespan Changes in Brain and Cognition, University of Oslo, Oslo, Norway, [6] Department of Clinical Sciences, Psychiatry, Umeå University, Umeå, Sweden, [7] Department of Medical Sciences, Neurology, Uppsala University, Uppsala, Sweden

Two of the main features of idiopathic normal pressure hydrocephalus (iNPH) are disturbed gait and cognition. These features are typically investigated separately, but here we combined walking with a cognitive task to investigate if older adults with iNPH were more susceptible to dual-task interference on walking than those without iNPH. In total, 95 individuals from the general population participated in our study. Of these, 20 were classified as Possible iNPH (median [interquartile range, IQR] 80 years [75–82.5]) and 75 as Unlikely iNPH (74 years [72–78]). Conversation, 10-m walking, semantic and phonemic verbal fluency were performed either combined or independently. "Stopping walking while talking" was noted. Pairwise comparisons and multiple logistic regression analyses were used. We found that the Possible iNPH group was older, stopped walking more frequently during the conversation, and had a slower single-task pace. The dual-task pace was slower for both groups. Only single-task walking pace could predict Possible iNPH when adjusted for age. We could establish a dual-task cost on gait performance in this sample of older adults from the general population, but the cost was not exclusive for individuals with Possible iNPH. To further assess the value of dual-task testing in iNPH, including observations of stopping walking while talking, a study of a clinical iNPH material with more severe symptoms would be valuable.

Keywords: dual-task, idiopathic normal pressure hydrocephalus, ageing, cognition, neuropsychology, older adults

OPEN ACCESS

Edited by:
Wai Haung (Ho) Yu,
University of Toronto, Canada

Reviewed by:
Ville Leinonen,
Kuopio University Hospital, Finland
Fanpei G. Yang,
National Tsing Hua University, Taiwan

****Correspondence:***
Otto Lilja-Lund
otto.lilja-lund@umu.se

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *25 March 2022*
Accepted: *27 April 2022*
Published: *30 May 2022*

Citation:
Lilja-Lund O, Nyberg L, Maripuu M and Laurell K (2022) Dual-Task Performance in Older Adults With and Without Idiopathic Normal Pressure Hydrocephalus. Front. Aging Neurosci. 14:904194. doi: 10.3389/fnagi.2022.904194

INTRODUCTION

Idiopathic normal pressure hydrocephalus (iNPH) is a disorder affecting older adults with the hallmark symptoms of gait disturbances, cognitive decline, and incontinence (Hakim and Adams, 1965; Mori et al., 2012). Radiological characteristics include disproportionately enlarged subarachnoid-space hydrocephalus (DESH) (Kitagaki et al., 1998), wide temporal horns (Lilja-Lund et al., 2020), and a narrow callosal angle (Virhammar et al., 2014). Untreated, iNPH reduces life

expectancy, autonomy, and health-related quality of life (Petersen et al., 2014; Andrén et al., 2020). Andersson et al. (2019) found a prevalence in the range of 1.5–3.7% in older adults, increasing with age to 8.9% among people aged 80 years and above. The cause of iNPH is still unknown, but candidate factors are disturbed cerebrospinal fluid (CSF) dynamics and vascular risk factors, such as diabetes mellitus and arterial hypertension (Bräutigam et al., 2019).

Correct diagnosis and treatment improve the individual's quality of life, as well as have socioeconomical gains (Petersen et al., 2014; Tullberg et al., 2018). The treatment of iNPH with CSF shunting can normalize mortality (Andrén et al., 2020), improve or maintain cognitive function over time (Kambara et al., 2021), decrease symptoms of gait and balance disturbances, and reduce incontinence (Todisco et al., 2020). Furthermore, iNPH is believed to be underdiagnosed, and increased awareness among clinicians, as well as the general public, could be an important step toward early recognition (Tullberg et al., 2018; Nakajima et al., 2021). Timely intervention is important as potential gains and improvement increase with early shunting (Andrén et al., 2014; Kambara et al., 2021). Currently, diagnosis can only be based on clinical examination and radiology, and new methods of assessment could contribute to the early detection and understanding of the disorder.

Dual-tasking is characterized by performing two separate tasks at the same time, straining limited cognitive resources (Pashler, 1994). An everyday example is walking and talking. Cognition and gait are negatively affected during dual-tasking, more so than what increased age alone represents (Verhaeghen et al., 2003; Al-Yahya et al., 2011). There are several ways to investigate dual-tasking, such as having a conversation while walking (Beauchet et al., 2009). Lundin-Olsson et al. (1997) found that older adults who stopped walking while talking (SWWT) had a significantly higher risk of falling within the next 6 months. This is important as fall accidents in older adults more frequently result in death or severe injuries compared with falls among adults 20–65 years old (Sibelius and Dahlström, 2005; Bridenbaugh and Kressig, 2015). Another method of testing dual-tasking is verbal fluency combined with walking (Al-Yahya et al., 2011).

Given that higher age, gait disturbances, and cognitive decline are characteristic features of iNPH, difficulties in "walking-and-talking" dual-tasking seem likely. However, there are only a few studies on dual-tasking in iNPH (Armand et al., 2011; Allali et al., 2013, 2017a,b; Schniepp et al., 2017; Selge et al., 2018). Dual-task testing has revealed positive tap test responders (Allali et al., 2013, 2017a), possibly better than single-task (Armand et al., 2011), and the effect was best 3 days post drainage (Schniepp et al., 2017). In addition, iNPH patients with apathy had an increased stride time variability during backward-counting and walking compared with iNPH without apathy (Allali et al., 2017b). Dual-task performance was also found to differentiate between progressive supranuclear palsy and iNPH, with better performance for the latter (Selge et al., 2018). The hypothesis here was that subjects with Possible iNPH are more susceptible to dual-task interference compared with older adults from the general population without iNPH. To the best of our knowledge, this hypothesis has not been tested before.

MATERIALS AND METHODS

Participants

The participants were recruited from an epidemiological study on iNPH among inhabitants of Jämtland Härjedalen, Sweden, aged 65 years or older (Kockum et al., 2018; Andersson et al., 2019; Lilja-Lund et al., 2020). The final sample consisted of 95 individuals with and without symptoms of iNPH. **Figure 1** shows the selection flowchart.

Method of Diagnosing

The Japanese guidelines, 2nd edition, were used to diagnose iNPH (Mori et al., 2012). The iNPH symptom scale was used to assess gait, balance, incontinence, and neuropsychology (Hellström et al., 2012). The participants underwent computed tomography (CT) of the brain (GE MD Optima CT540). The protocol used was 120 kV, 300 MaS, rotation time 0.5 s with a pitch of 0.6, generating a slice thickness of 0.6 mm with 4 mm reconstructions in three planes. Radiological markers of iNPH were scored using the iNPH Radscale (Kockum et al., 2018, 2020). Cognitive status was screened using the Mini-mental State Examination (MMSE) (Folstein et al., 1975). Symptoms of depression were screened using the Geriatric Depression Scale (GDS-15) (Kørner et al., 2006).

Single-and Dual-Tasks

The participants were asked about how they got to the hospital on the way to the examination room from the waiting room.

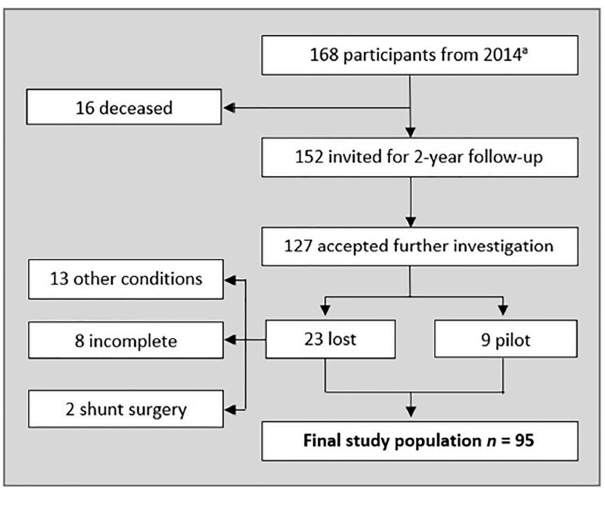


FIGURE 1 | Flow-chart representing included participants. [a] see Andersson et al., 2019. Other conditions severely affecting gait and/or cognition was excluded from the sample (Alzheimer's disease, hip-surgery, cancer, visual impairment, spinal stenosis, and secondary hydrocephalus. Five of them failed the dual task, and three used walking aid). Three declined neurological examination, three had incomplete neuropsychological tests, and two declined imaging. Two participants had been given a shunt after the 2014 study. Nine randomly picked participants (all diagnosed with Unlikely iNPH) were given a pilot test protocol. The number of pilots needed was based on when the test protocol was fulfilling the aim of the study.

Observations on whether they stopped walking during this conversation was noted. Similarly, SWWT was registered during the dual-task testing described below.

The single-tasks consisted of timed silent 10-m walking at a normal pace and verbal fluency while seated. The pace was timed with a stopwatch. The participants were asked to "name as many things as possible you can think of that you can find in a grocery store" during the semantic test. On the phonemic test, they were asked "to say as many words as possible that you can think of starting with the letter A, except numbers, names of persons, or places." The time-limit on single-task fluency was 60 s.

The dual-task testing consisted of walking 10-m at a preferred pace while performing a semantic or phonological verbal fluency test. Things you can find in a home were used as a semantic fluency task and the letter F on the phonemic fluency. The investigator counted the words using a mechanical tally counter. Repeated words were excluded from all fluency counts.

Statistical Analyses

Statistical analyses were conducted using IBM SPSS Statistics 27 (IBM Corp., Armonk, NY, United States). The assumption of normal distribution was investigated and tested with the Shapiro–Wilk test. Non-parametric tests were chosen based on the group size, skewed distribution, and types of variables analyzed. Distribution of sex and frequency of SWWT were tested using chi-square and Fisher's exact tests. Differences between groups were tested using the Mann–Whitney *U*-tests. Multiple logistic regression analyses were performed to adjust for age. Differences between tasks were tested using the Friedman test and Wilcoxon signed ranks tests. The following equation was used to calculate the dual-task cost:

$$\frac{DTtime - STtime}{STtime} * 100$$

The level of significance was set to $p < 0.05$ with Bonferroni correction applied when appropriate.

RESULTS

The final sample consisted of 95 participants, with 75 in the Unlikely iNPH group and 20 in the Possible iNPH group. The Possible group was older (median [interquartile range, IQR]; 80 years [75–82.5] vs. 74 years [72–78], $p = 0.004$), had more symptoms (iNPH score 73.9 [67–81] vs. 90.5 [80.7–95.3], $p < 0.001$), and more radiological signs of iNPH (Radscale score 4 [3–5] vs. 2 [1–3], $p < 0.001$). The two groups did not differ in sex, education, MMSE, or GDS-15 (n.s.), as shown in **Table 1**.

Observations of SWWT during conversation revealed that two (10%) from the Possible iNPH group stopped walking and none (0%) in the Unlikely group ($p = 0.044$). During dual-task testing, the number of people who stopped walking was similar for both groups (n.s.), as shown in **Table 2**.

Both groups reduced their pace during dual-tasking when compared to single-task walking [Possible iNPH χ^2 (2) = 26.7, $p < 0.001$; and Unlikely iNPH χ^2 (2) = 103.9, $p < 0.001$]. The *post-hoc* analysis showed that the pace was reduced during semantic ($z = -3.1$, $p = 0.002$) and phonemic ($z = -3.9$, $p < 0.001$) dual-task fluency for the Possible iNPH group. Equally, the pace slowed down during semantic ($z = -7.1$, $p < 0.001$) and phonemic ($z = -7.4$, $p < 0.001$) dual-task fluency for the Unlikely iNPH group. The dual-task cost on pace did not differ between the two groups. The increase in walking time during semantic fluency for the Possible iNPH group was (median, IQR) 38% (5–61%), and 34% (18–71%) for the Unlikely iNPH (n.s.). During phonemic fluency, the increase in walking time for the Possible iNPH was 41% (25–62%), and 45% (21–112%) for the Unlikely iNPH (n.s.), as shown in **Table 3**.

There were some differences between the groups in single-task performance. The Possible iNPH group had a slower single-task pace ($p < 0.001$) and produced fewer words during semantic single-task fluency ($p = 0.024$) but generated a similar number of words during phonemic single-task fluency (n.s.) compared with the Unlikely iNPH group (as shown in **Table 3**). Single-task pace predicted Possible iNPH correctly (odds ratio [*OR*] [95% *CI*] 1.45 [1.02–2.06], $p = 0.038$) when adjusting for age. However, semantic single-task fluency adjusted for age could not predict iNPH diagnosis correctly (n.s.).

Performance was worse for both groups on the phonemic tasks compared with the semantic tasks. The dual-task pace was slower ($z = -2.5$, $p = 0.013$), and word production decreased during single-task ($z = -3.9$, $p < 0.001$) and dual-task phonemic fluency ($z = -3.4$, $p < 0.001$) for the Possible iNPH group. Similarly, dual-task pace ($z = -4.0$, $p < 0.001$), and single-task ($z = -7.5$, $p < 0.001$) and dual-task fluency ($z = -6.2$, $p < 0.001$) decreased for the Unlikely iNPH group, as shown in **Table 3**.

TABLE 1 | Descriptive statistics of participants.

	Possible iNPH	Unlikely iNPH	χ^2	*p*
Participants *n* (female%)	20 (45%)	75 (59%)	1.196	0.317
	Md (IQR)	Md (IQR)	*U*	
Age (years)	80 (75–82.5)	74 (72–78)	438	**0.004**
INPH symptom scale[a] (0–100)	73.9 (67–81)	90.5 (80.7–95.3)	264	**<0.001**
INPH Radscale[b] (0–12)	4 (3–5)	2 (1–3)	254.5	**<0.001**
Education (years)	9 (7–12.5)	9 (7–13)	684	0.541
MMSE[c] (0–30)	27 (26–28)	27 (26–28)	712.5	0.795
GDS-15[d] (0–15)	2 (0.25–3)	1 (0–2)	550	0.059

Md, median; IQR, interquartile range. [a]Hellström et al. (2012); higher score = less symptoms. [b]Kockum et al. (2018); higher score = more symptoms. [c]Folstein et al. (1975); higher score = less symptoms. [d]Kørner et al. (2006); higher score = more symptoms. Significant values of p are in bold.

TABLE 2 | Frequency of participants who stopped walking while talking (SWWT).

	Possible iNPH *n* = 20		Unlikely iNPH *n* = 75		*Fischer's exact test p*-value
	Stops *n* (%)	Walks *n* (%)	Stops *n* (%)	Walks *n* (%)	
Conversation	2 (10)	18 (90)	0 (0)	73 (100)	**0.044**
Semantic fluency	3 (15)	17 (85)	7 (9)	68 (91)	0.434
Phonemic fluency	3 (15)	17 (85)	18 (24)	57 (76)	0.548

Significant values of p are in bold.

TABLE 3 | Mann-Whitney *U*-tests of group differences in walking speed and verbal fluency.

	Possible iNPH *n* = 20	Unlikely iNPH *n* = 75		
	Md (IQR) [min–max]	**Md (IQR) [min–max]**	***U***	***p***
Single-task				
10-meter walking (sec.)	10.1 (9.6–11.7) [7–31]	8.8 (7.6–9.8) [6–13]	337	**<0.001**
Semantic fluency[a]	17.5 (12.5–21.8) [11–30]	22 (17–27) [8–44]	502.5	**0.024**
Phonemic fluency[a]	8.5 (6–11.8) [2–15]	10 (7–13) [1–23]	648.5	0.352
Dual-task				
Semantic 10-m (sec.)	14.0 (12.3–17.0) [8–33]	12.0 (9.0–17.0) [6–41]	555	0.074
Phonemic 10-m (sec.)	16.0 (14.0–17.0) [10–98]	13.0 (10.0–20.0) [7–43]	557	0.077
Semantic fluency[a]	8 (6–10.8) [4–15]	8 (6–10) [2–17]	735.5	0.894
Phonemic fluency[a]	5.5 (4–6.8) [1–16]	5 (4–6) [1–17]	661.5	0.414

Md, median; IQR, interquartile range. Significant values of p are in bold. [a]*Number of words.*

DISCUSSION

We hypothesized that older adults with Possible iNPH could be more susceptible for dual-task interference. We found that the Possible iNPH group SWWT more frequently during the conversation compared with the Unlikely iNPH group. Furthermore, the Possible iNPH group was slower during single-task walking compared with the Unlikely iNPH group. However, in contrast to our hypothesis, no difference in SWWT was found between the groups during the semantic or phonemic dual-task testing and the dual-task cost on pace and word production was similar for both groups.

The need to stop walking to talk during the normal conversation did not occur for the Unlikely group, but it was 10% in the Possible iNPH group. In their study, Lundin-Olsson et al. (1997) presumed that stopping walking happens due to attentional constraints when performing two tasks at once but no cognitive assessments beyond the MMSE score for the whole sample were reported. However, this assumption has support in other studies confirming that executive functioning in older adult fallers was reduced compared with non-fallers and younger adults and that the risk of falling increases with poor cognitive performance (Springer et al., 2006; Montero-Odasso et al., 2012). Montero-Odasso et al. (2012) argue in their review that gait and cognition should be viewed as intertwined aspects of aging and not as separate domains. Proactive measures against falling (such as exercise, revision of medications, and environmental factors) should be taken since older adults who stop talking is at greater risk of cognitive decline and falling (American Geriatrics Society et al., 2001).

Another indication of an increased cognitive load during dual-tasking was that both groups decreased their pace during dual-task fluency, in line with the "bottleneck" view of available executive capacity (Yogev-Seligmann et al., 2008). Other studies have found that verbal fluency had a negative impact on walking speed, and there is evidence of left prefrontal cortical (PFC) engagement relating to gait speed control, regardless of global physical strain (Harada et al., 2009; Al-Yahya et al., 2011). The limitations of a central executive coordinating multi-sensor processing in a top-down fashion involving the PFC could be related to the reduction of pace during dual-tasking (Katus and Eimer, 2019). One study found that young adults had increased PFC activity while talking during walking compared with an older population, and the authors argue that this could be caused by the age-related decrease in PFC functioning found in older age (Holtzer et al., 2011).

We assumed that dual-tasking would increase the cognitive load and potentially amplify discrepancies between the two groups in our study. However, in contrast to the difference in SWWT during normal conversation, the two groups did not differ in the frequency of SWWT during the dual-task fluency tests. One possible explanation is that normal conversation is less taxing compared with the verbal fluency tasks and therefore less likely to trigger SWWT in the Unlikely group. The increase in the number of participants who SWWT during the verbal fluency tasks supports this interpretation. We could not compare potential dual-task cost to walking speed during conversation as this was not noted.

Interestingly, walking speed during dual-tasking did not statistically differ between the two groups, but a clear dual-task effect emerged in both groups. One effect was the dispersion of walking time within the groups. During single-task walking, the IQR in speed was approximately 2 s for both groups. During dual-tasking, the widest IQR for the Possible iNPH was 5 s during semantic fluency and 10 s for the Unlikely iNPH during phonemic fluency. Hence, even though the Possible iNPH median walking time repeatedly surpassed the Unlikely group by approximately 2 s, the heterogeneous dual-task effect in the Unlikely group contributed to the value of *p* exceeding the alpha level. In summary, we could establish a dual-task cost, but the effects were not exclusive to the Possible iNPH, and individual differences were marked.

There were also differences regarding the performance of single-task. Gait disturbances are a core feature in iNPH and the Possible iNPH group had a slower single-task pace when adjusting for age. Hence, testing single-task walking is

fundamental in investigating iNPH. Moreover, the Possible iNPH group had a lower word production during single-task semantic fluency compared with the Unlikely iNPH group. However, adjusting for age revealed that the higher age significantly contributed more to the classification of diagnostic categories. Other studies have found that increased age was associated with a reduction in verbal fluency (Tombaugh et al., 1999). However, it is important to consider that age does not completely exclude effects following iNPH as the prevalence of the disorder increases with age (Andersson et al., 2019). The study sample comes from a population-based epidemiological study, and the higher age for the Possible iNPH group is in accordance with the literature (Iseki et al., 2014; Jaraj et al., 2014; Nakajima et al., 2021). Other demographic variables did not differ between the groups (the sex distribution, level of education, MMSE score, and GDS-15 score).

Previous studies have shown that phonemic fluency is more related to the PFC, and semantic fluency to the hippocampi (Glikmann-Johnston et al., 2015). Inferior performance on a more hippocampi-dependent task in iNPH is intriguing as widening of the temporal horns surrounding the hippocampi has been associated with all main symptoms of iNPH (Lilja-Lund et al., 2020). Furthermore, a study using verbal fluency during walking to evaluate the effects of tap-testing in iNPH revealed improvements in semantic fluency post-tap but not in phonemic fluency, possibly indicating that semantic fluency is sensitive to iNPH (Allali et al., 2017a).

Studies comparing patients with iNPH and clinical populations found that iNPH performed worse on phonemic fluency compared with Alzheimer's disease and *de novo* Parkinson's disease, but not on semantic fluency (Miyoshi et al., 2005; Picascia et al., 2019). However, it is important to note that our participants were relatively healthy. The patients with iNPH in Picascia et al. (2019) had more severe and diffuse cognitive symptoms compared with the *de novo* Parkinson's disease; a group with less severe cognitive decline. The cognitive state was inferior for all participants in Miyoshi et al. (2005) compared with our sample. Most studies on iNPH are retrospective studies, often related to shunting. Our sample from the general population had less progressed symptoms and represents what clinicians might meet in their practice, before diagnosis. On the other hand, this likely contributed to the relatively few differences between the groups in our study.

The participants were not instructed to prioritize any of the tasks (pace or fluency) during the dual-task testing. In hindsight and future studies on dual-tasking, it would be interesting to question the participants afterward if they prioritized any of the tasks intentionally, even though they were not instructed to do so. Still, 22% of the participants stopped walking during the most demanding task (phonemic dual-task) in violation of the instructions. Stopping walking and talking can cause extreme outliers in time. Future studies could investigate if this standstill is more frequent in specific groups, e.g., more severe iNPH, Alzheimer's disease, and notably Parkinson's disease. A recent meta-analysis reviewed the effects of dual-tasking on Parkinson's disease with the overall conclusion that it affects walking speed negatively, however, they did not mention SWWT or "freezing of gait" (Raffegeau et al., 2019). It would be interesting to include dual-task effects beyond pace when studying neurocognitive disorders.

"Stops walking while talking" has been suggested as a clinical test of dual-tasking but a lack of standardized questions has raised some concerns (Yogev-Seligmann et al., 2008; Beauchet et al., 2009). A strength of our study was that we had the same question and conditions when testing SWWT. There are some limitations to our study as well. We had to exclude direct comparison of dual-task effects on verbal fluency since different letters and categories can be confounding when comparing fluency and using the same items would introduce more learning effects (Tombaugh et al., 1999; Lezak et al., 2012). Video recordings of walking tests would have been helpful in additional analyses. The choice to exclude invasive tests, such as CSF analysis, was made to minimize associated risks and drop-out. However, using clinical and radiological features to diagnose iNPH is viable in epidemiological studies on iNPH (Mori et al., 2012).

Future research could investigate if SWWT can help assess and predict shunt outcomes. Studies on SWWT typically focus on fall-incidents (Ayers et al., 2014). It would be interesting to evaluate the predictive value of SWWT in diagnosing iNPH or other neurocognitive disorders in longitudinal studies. The topic of the conversation in our study (how they got to the clinic) prompted the episodic memory of a recent event involving spatial navigation. It would be interesting to investigate if the topic for the conversation matters for dual-task cost or SWWT, for example, whether there is a difference between semantic memory vs. episodic memory. Given that semantic and episodic memory display different trajectories of change across the lifespan, targeting different memory systems could be of interest (Nyberg et al., 2003, 2020).

CONCLUSION

The Possible iNPH group SWWT more frequently during conversation and had a slower single-task walking time. There was a distinct dual-task interference with walking speed being negatively affected during verbal fluency, but the dual-task cost was similar in both groups. The use of dual-tasking needs to be further investigated to delineate its usefulness in iNPH. It is noteworthy that dual-task effects can alter performance in unexpected ways compared with standard single-task testing.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Regional Ethical Review Board in Umeå (Dnr 2014/180-31 and Dnr 2017-167-32M) and the Radiation

Protection Committee (2014-10-03 and 2017-04-24). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

OL-L conducted the neuropsychological testing and wrote the first draft of the manuscript. KL performed the neurological examinations and decided the diagnosis. All authors were involved in the design, statistical analysis, revision of the text, and approved the submitted version.

FUNDING

This work was funded by Umeå University, Uppsala University, and the County Council of Jämtland Härjedalen. LN was supported by a Scholar grant from the Knut and Alice Wallenberg (KAW) Foundation.

ACKNOWLEDGMENTS

We are grateful for the imaging by Karin Kockum and the assistance of Johanna Andersson.

REFERENCES

Allali, G., Laidet, M., Armand, S., Momjian, S., Marques, B., Saj, A., et al. (2017a). A combined cognitive and gait quantification to identify normal pressure hydrocephalus from its mimics: the Geneva's protocol. *Clin. Neurol. Neurosurg.* 160, 5–11. doi: 10.1016/j.clineuro.2017.06.001

Allali, G., Laidet, M., Armand, S., Saj, A., Krack, P., and Assal, F. (2017b). Apathy and higher level of gait control in normal pressure hydrocephalus. *Int. J. Psychophysiol.* 119, 127–131. doi: 10.1016/j.ijpsycho.2016.12.002

Allali, G., Laidet, M., Beauchet, O., Herrmann, F. R., Assal, F., and Armand, S. (2013). Dual-task related gait changes after CSF tapping: a new way to identify idiopathic normal pressure hydrocephalus. *J. NeuroEngineering Rehabil.* 10:117. doi: 10.1186/1743-0003-10-117

Al-Yahya, E., Dawes, H., Smith, L., Dennis, A., Howells, K., and Cockburn, J. (2011). Cognitive motor interference while walking: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 35, 715–728. doi: 10.1016/j.neubiorev.2010.08.008

American Geriatrics Society, British Geriatrics Society, and American Academy Of Orthopaedic Surgeons Panel On Falls Prevention (2001). Guideline for the prevention of falls in older persons. *J. Am. Geriatr. Soc.* 49, 664–672. doi: 10.1046/j.1532-5415.2001.49115.x

Andersson, J., Rosell, M., Kockum, K., Lilja-Lund, O., Söderström, L., and Laurell, K. (2019). Prevalence of idiopathic normal pressure hydrocephalus: a prospective, population-based study. *PLoS One* 14:e0217705. doi: 10.1371/journal.pone.0217705

Andrén, K., Wikkelsø, C., Sundström, N., Israelsson, H., Agerskov, S., Laurell, K., et al. (2020). Survival in treated idiopathic normal pressure hydrocephalus. *J. Neurol.* 267, 640–648. doi: 10.1007/s00415-019-09598-1

Andrén, K., Wikkelsø, C., Tisell, M., and Hellström, P. (2014). Natural course of idiopathic normal pressure hydrocephalus. *J. Neurol. Neurosurg. Psychiatry* 85, 806–810. doi: 10.1136/jnnp-2013-306117

Armand, S., Allet, L., Landis, T., Beauchet, O., Assal, F., and Allali, G. (2011). Interest of dual-task-related gait changes in idiopathic normal pressure hydrocephalus: dual-task-related gait changes and idiopathic normal pressure hydrocephalus. *Eur. J. Neurol.* 18, 1081–1084. doi: 10.1111/j.1468-1331.2010.03242.x

Ayers, E. I., Tow, A. C., Holtzer, R., and Verghese, J. (2014). Walking while talking and falls in aging. *Gerontology* 60, 108–113. doi: 10.1159/000355119

Beauchet, O., Annweiler, C., Dubost, V., Allali, G., Kressig, R. W., Bridenbaugh, S., et al. (2009). Stops walking when talking: a predictor of falls in older adults? *Eur. J. Neurol.* 16, 786–795. doi: 10.1111/j.1468-1331.2009.02612.x

Bräutigam, K., Vakis, A., and Tsitsipanis, C. (2019). Pathogenesis of idiopathic normal pressure hydrocephalus: a review of knowledge. *J. Clin. Neurosci.* 61, 10–13. doi: 10.1016/j.jocn.2018.10.147

Bridenbaugh, S. A., and Kressig, R. W. (2015). Motor cognitive dual tasking. Early detection of gait impairment, fall risk and cognitive decline. *Z. Für Gerontol. Geriatr.* 48, 15–21. doi: 10.1007/s00391-014-0845-0

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Glikmann-Johnston, Y., Oren, N., Hendler, T., and Shapira-Lichter, I. (2015). Distinct functional connectivity of the hippocampus during semantic and phonemic fluency. *Neuropsychologia* 69, 39–49. doi: 10.1016/j.neuropsychologia.2015.01.031

Hakim, S., and Adams, R. D. (1965). The special clinical problem of symptomatic hydrocephalus with normal cerebrospinal fluid pressure: observations on cerebrospinal fluid hydrodynamics. *J. Neurol. Sci.* 2, 307–327. doi: 10.1016/0022-510X(65)90016-X

Harada, T., Miyai, I., Suzuki, M., and Kubota, K. (2009). Gait capacity affects cortical activation patterns related to speed control in the elderly. *Exp. Brain Res.* 193, 445–454. doi: 10.1007/s00221-008-1643-y

Hellström, P., Klinge, P., Tans, J., and Wikkelsø, C. (2012). A new scale for assessment of severity and outcome in iNPH. *Acta Neurol. Scand.* 126, 229–237. doi: 10.1111/j.1600-0404.2012.01677.x

Holtzer, R., Mahoney, J. R., Izzetoglu, M., Izzetoglu, K., Onaral, B., and Verghese, J. (2011). fNIRS study of walking and walking while talking in young and old individuals. *J. Gerontol. Ser. A* 66A, 879–887. doi: 10.1093/gerona/glr068

Iseki, C., Takahashi, Y., Wada, M., Kawanami, T., Adachi, M., and Kato, T. (2014). Incidence of idiopathic normal pressure hydrocephalus (iNPH): a 10-year follow-up study of a rural community in Japan. *J. Neurol. Sci.* 339, 108–112. doi: 10.1016/j.jns.2014.01.033

Jaraj, D., Rabiei, K., Marlow, T., Jensen, C., Skoog, I., and Wikkelsø, C. (2014). Prevalence of idiopathic normal-pressure hydrocephalus. *Neurology* 82, 1449–1454. doi: 10.1212/WNL.0000000000000342

Kambara, A., Kajimoto, Y., Yagi, R., Ikeda, N., Furuse, M., Nonoguchi, N., et al. (2021). Long-term prognosis of cognitive function in patients with idiopathic normal pressure hydrocephalus after shunt surgery. *Front. Aging Neurosci.* 12:617150. doi: 10.3389/fnagi.2020.617150

Katus, T., and Eimer, M. (2019). The sources of dual-task costs in multisensory working memory tasks. *J. Cogn. Neurosci.* 31, 175–185. doi: 10.1162/jocn_a_01348

Kitagaki, H., Mori, E., Ishii, K., Yamaji, S., Hirono, N., and Imamura, T. (1998). CSF spaces in idiopathic normal pressure hydrocephalus: morphology and volumetry. *Am. J. Neuroradiol.* 19, 1277–1284.

Kockum, K., Lilja-Lund, O., Larsson, E.-M., Rosell, M., Söderström, L., Virhammar, J., et al. (2018). The idiopathic normal-pressure hydrocephalus radscale: a radiological scale for structured evaluation. *Eur. J. Neurol.* 25, 569–576. doi: 10.1111/ene.13555

Kockum, K., Virhammar, J., Riklund, K., Söderström, L., Larsson, E.-M., and Laurell, K. (2020). Diagnostic accuracy of the iNPH Radscale in idiopathic normal pressure hydrocephalus. *PLoS One* 15:e0232275. doi: 10.1371/journal.pone.0232275

Kørner, A., Lauritzen, L., Abelskov, K., Gulmann, N., Marie, A., Wedervang-Jensen, T., et al. (2006). The Geriatric Depression Scale and the Cornell Scale for Depression in Dementia. A validity study. *Nord J Psychiatry* 60, 360–364. doi: 10.1080/08039480600937066

Lezak, M. D., Howieson, D. B., and Loring, D. W. (2012). *Neuropsychological Assessment*, 5th Edn. New York, NY: Oxford University Press.

Lilja-Lund, O., Kockum, K., Hellström, P., Söderström, L., Nyberg, L., and Laurell, K. (2020). Wide temporal horns are associated with cognitive dysfunction, as

well as impaired gait and incontinence. *Sci. Rep.* 10:18203. doi: 10.1038/s41598-020-75381-2

Lundin-Olsson, L., Nyberg, L., and Gustafson, Y. (1997). "Stops walking when talking" as a predictor of falls in elderly people. *Lancet* 349:617. doi: 10.1016/S0140-6736(97)24009-2

Miyoshi, N., Kazui, H., Ogino, A., Ishikawa, M., Miyake, H., Tokunaga, H., et al. (2005). Association between cognitive impairment and gait disturbance in patients with idiopathic normal pressure hydrocephalus. *Dement. Geriatr. Cogn. Disord.* 20, 71–76. doi: 10.1159/000085858

Montero-Odasso, M., Verghese, J., Beauchet, O., and Hausdorff, J. M. (2012). Gait and cognition: a complementary approach to understanding brain function and the risk of falling. *J. Am. Geriatr. Soc.* 60, 2127–2136. doi: 10.1111/j.1532-5415.2012.04209.x

Mori, E., Ishikawa, M., Kato, T., Kazui, H., Miyake, H., Miyajima, M., et al. (2012). Guidelines for management of idiopathic normal pressure hydrocephalus: second edition. *Neurol. Med. Chir.* 52, 775–809. doi: 10.2176/nmc.52.775

Nakajima, M., Yamada, S., Miyajima, M., Ishii, K., Kuriyama, N., Kazui, H., et al. (2021). Guidelines for management of idiopathic normal pressure hydrocephalus (third edition): endorsed by the Japanese society of normal pressure hydrocephalus. *Neurol. Med. Chir.* 61, 63–97. doi: 10.2176/nmc.st.2020-0292

Nyberg, L., Boraxbekk, C.-J., Sörman, D. E., Hansson, P., Herlitz, A., Kauppi, K., et al. (2020). Biological and environmental predictors of heterogeneity in neurocognitive ageing. *Ageing Res. Rev.* 64:101184. doi: 10.1016/j.arr.2020.101184

Nyberg, L., Maitland, S. B., Rönnlund, M., Bäckman, L., Dixon, R. A., Wahlin, Å, et al. (2003). Selective adult age differences in an age-invariant multifactor model of declarative memory. *Psychol. Aging* 18, 149–160. doi: 10.1037/0882-7974.18.1.149

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychol. Bull.* 116, 220–224. doi: 10.1037/0033-2909.116.2.220

Petersen, J., Hellström, P., Wikkelsø, C., and Lundgren-Nilsson, Å (2014). Improvement in social function and health-related quality of life after shunt surgery for idiopathic normal-pressure hydrocephalus. *J. Neurosurg.* 121, 776–784. doi: 10.3171/2014.6.JNS132003

Picascia, M., Pozzi, N. G., Todisco, M., Minafra, B., Sinforiani, E., Zangaglia, R., et al. (2019). Cognitive disorders in normal pressure hydrocephalus with initial parkinsonism in comparison with de novo Parkinson's disease. *Eur. J. Neurol.* 26, 74–79. doi: 10.1111/ene.13766

Raffegeau, T. E., Krehbiel, L. M., Kang, N., Thijs, F. J., Altmann, L. J. P., Cauraugh, J. H., et al. (2019). A meta-analysis: Parkinson's disease and dual-task walking. *Parkinsonism Relat. Disord.* 62, 28–35. doi: 10.1016/j.parkreldis.2018.12.012

Schniepp, R., Trabold, R., Romagna, A., Akrami, F., Hesselbarth, K., Wuehr, M., et al. (2017). Walking assessment after lumbar puncture in normal-pressure hydrocephalus: a delayed improvement over 3 days. *J. Neurosurg.* 126, 148–157. doi: 10.3171/2015.12.JNS151663

Selge, C., Schoeberl, F., Zwergal, A., Nuebling, G., Brandt, T., Dieterich, M., et al. (2018). Gait analysis in PSP and NPH: dual-task conditions make the difference. *Neurology* 90, e1021–e1028. doi: 10.1212/WNL.0000000000005168

Sibelius, J., and Dahlström, F. (2005). *Samhällets Kostnader för Fallolyckor*. Stockholm: Myndigheten för samhällsskydd och beredskap.

Springer, S., Giladi, N., Peretz, C., Yogev, G., Simon, E. S., and Hausdorff, J. M. (2006). Dual-tasking effects on gait variability: the role of aging, falls, and executive function. *Mov. Disord.* 21, 950–957. doi: 10.1002/mds.20848

Todisco, M., Picascia, M., Pisano, P., Zangaglia, R., Minafra, B., Vitali, P., et al. (2020). Lumboperitoneal shunt in idiopathic normal pressure hydrocephalus: a prospective controlled study. *J. Neurol.* 267, 2556–2566. doi: 10.1007/s00415-020-09844-x

Tombaugh, T. N., Kozak, J., and Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Arch. Clin. Neuropsychol.* 14, 167–177. doi: 10.1016/S0887-6177(97)00095-4

Tullberg, M., Persson, J., Petersen, J., Hellström, P., Wikkelsø, C., and Lundgren-Nilsson, Å (2018). Shunt surgery in idiopathic normal pressure hydrocephalus is cost-effective—a cost utility analysis. *Acta Neurochir.* 160, 509–518. doi: 10.1007/s00701-017-3394-7

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., and Cerella, J. (2003). Aging and dual-task performance: a meta-analysis. *Psychol. Aging* 18, 443–460. doi: 10.1037/0882-7974.18.3.443

Virhammar, J., Laurell, K., Cesarini, K. G., and Larsson, E.-M. (2014). The callosal angle measured on MRI as a predictor of outcome in idiopathic normal-pressure hydrocephalus. *J. Neurosurg.* 120, 178–184. doi: 10.3171/2013.8.JNS13575

Yogev-Seligmann, G., Hausdorff, J. M., and Giladi, N. (2008). The role of executive function and attention in gait. *Mov. Disord.* 23, 329–342. doi: 10.1002/mds.21720

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fnagi.2022.857521

Comparison of Conventional Logistic Regression and Machine Learning Methods for Predicting Delayed Cerebral Ischemia After Aneurysmal Subarachnoid Hemorrhage: A Multicentric Observational Cohort Study

Ping Hu[1,2†], Yuntao Li[1†], Yangfan Liu[2†], Geng Guo[3†], Xu Gao[4†], Zhongzhou Su[5†], Long Wang[1], Gang Deng[1], Shuang Yang[6,7], Yangzhi Qi[1], Yang Xu[1], Liguo Ye[1], Qian Sun[1], Xiaohu Nie[5], Yanqi Sun[3], Mingchang Li[1], Hongbo Zhang[8] and Qianxue Chen[1*]*

[1] Department of Neurosurgery, Renmin Hospital of Wuhan University, Wuhan, China, [2] Department of Neurosurgery, Affiliated Hospital of Panzhihua University, Panzhihua, China, [3] Department of Neurosurgery, First Hospital of Shanxi Medical University, Taiyuan, China, [4] Department of Neurosurgery, General Hospital of Northern Theater Command, Shenyang, China, [5] Department of Neurosurgery, Huzhou Central Hospital, Huzhou, China, [6] School of Physics and Technology, Wuhan University, Wuhan, China, [7] School of Electronic Information and Automation, Guilin University of Aerospace Technology, Guilin, China, [8] Department of Neurosurgery, The Second Affiliated Hospital of Nanchang University, Nanchang, China

Background: Timely and accurate prediction of delayed cerebral ischemia is critical for improving the prognosis of patients with aneurysmal subarachnoid hemorrhage. Machine learning (ML) algorithms are increasingly regarded as having a higher prediction power than conventional logistic regression (LR). This study aims to construct LR and ML models and compare their prediction power on delayed cerebral ischemia (DCI) after aneurysmal subarachnoid hemorrhage (aSAH).

Methods: This was a multicenter, retrospective, observational cohort study that enrolled patients with aneurysmal subarachnoid hemorrhage from five hospitals in China. A total of 404 aSAH patients were prospectively enrolled. We randomly divided the patients into training ($N = 303$) and validation cohorts ($N = 101$) according to a ratio of 75–25%. One LR and six popular ML algorithms were used to construct models. The area under the receiver operating characteristic curve (AUC), accuracy, balanced accuracy, confusion matrix, sensitivity, specificity, calibration curve, and Hosmer–Lemeshow test were used to assess and compare the model performance. Finally, we calculated each feature of importance.

Results: A total of 112 (27.7%) patients developed DCI. Our results showed that conventional LR with an AUC value of 0.824 (95%CI: 0.73–0.91) in the validation cohort outperformed k-nearest neighbor, decision tree, support vector machine, and extreme gradient boosting model with the AUCs of 0.792 (95%CI: 0.68–0.9, $P = 0.46$), 0.675 (95%CI: 0.56–0.79, $P < 0.01$), 0.677 (95%CI: 0.57–0.77, $P < 0.01$), and 0.78 (95%CI:

OPEN ACCESS

Edited by:
S. Abid Hussaini, Columbia University Irving Medical Center, United States

Reviewed by:
Murad Megjhani, Columbia University Irving Medical Center, United States
Dagmar Verbaan, Academic Medical Center (AMC), Netherlands

***Correspondence:**
Hongbo Zhang
hongbozhang99@163.com
Qianxue Chen
chenqx666@whu.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Neuroinflammation and Neuropathy, a section of the journal Frontiers in Aging Neuroscience

Received: *18 January 2022*
Accepted: *09 May 2022*
Published: *17 June 2022*

Citation:
Hu P, Li Y, Liu Y, Guo G, Gao X, Su Z, Wang L, Deng G, Yang S, Qi Y, Xu Y, Ye L, Sun Q, Nie X, Sun Y, Li M, Zhang H and Chen Q (2022) Comparison of Conventional Logistic Regression and Machine Learning Methods for Predicting Delayed Cerebral Ischemia After Aneurysmal Subarachnoid Hemorrhage: A Multicentric Observational Cohort Study. Front. Aging Neurosci. 14:857521. doi: 10.3389/fnagi.2022.857521

0.68–0.87, $P = 0.50$). However, random forest (RF) and artificial neural network model with the same AUC (0.858, 95%CI: 0.78–0.93, $P = 0.26$) were better than the LR. The accuracy and the balanced accuracy of the RF were 20.8% and 11% higher than the latter, and the RF also showed good calibration in the validation cohort (Hosmer-Lemeshow: $P = 0.203$). We found that the CT value of subarachnoid hemorrhage, WBC count, neutrophil count, CT value of cerebral edema, and monocyte count were the five most important features for DCI prediction in the RF model. We then developed an online prediction tool (https://dynamic-nomogram.shinyapps.io/DynNomapp-DCI/) based on important features to calculate DCI risk precisely.

Conclusions: In this multicenter study, we found that several ML methods, particularly RF, outperformed conventional LR. Furthermore, an online prediction tool based on the RF model was developed to identify patients at high risk for DCI after SAH and facilitate timely interventions.

Clinical Trial Registration: http://www.chictr.org.cn, Unique identifier: ChiCTR2100044448.

Keywords: logistic regression, prediction model, delayed cerebral ischemia, subarachnoid hemorrhage, inflammatory response, machine learning

INTRODUCTION

Aneurysmal subarachnoid hemorrhage (aSAH) is a severe acute cerebrovascular disorder resulting in high morbidity and mortality; roughly 50% of aSAH survivors have permanent neurological deficits (Molyneux et al., 2005; Fugate and Rabinstein, 2012). Delayed cerebral ischemia (DCI) is the most frequent complication after aSAH, affecting ∼ 30% of patients, often causing serious damage because of its late diagnosis (Macdonald, 2014; Francoeur and Mayer, 2016). Hence, timely and accurate prediction of DCI is critical for the treatment and prognosis of patients with aSAH. A precise, reliable model for early prediction of DCI development is urgently needed.

Traditional logistic regression (LR) is the primary method to construct models for predicting disease outcomes. However, when LR is used for complex multivariate non-linear relationships, complex transformations are often required owing to low robustness and multicollinearity between variables (Tu, 1996). Machine learning (ML) is valuable for analyzing clinical data because it can fully employ input features and predict outcomes more accurately (Jordan and Mitchell, 2015). Several studies suggested that in DCI, ML models utilizing admission clinical characteristics have better predictive power than LR (Ramos et al., 2019; de Jong et al., 2021; Savarraj et al., 2021). However, the model performance is not generally high due to the incomplete clinical features. Admission clinical characteristics include baseline information, laboratory test results, and imaging data, and the fragmented application of these data may reduce predictive performance; therefore, these features must be systematically utilized. To the best of our knowledge, there is no study that utilizes relatively complete clinical features to construct ML and LR models, some of which were not compared in previous studies.

We determined several types of the currently most popular ML algorithms to achieve the following aims. First, we constructed and validated a conventional LR and several ML models based on relatively complete clinical features on admission. Second, we compared the predictive performances of the LR and ML models. Third, we established an online prediction tool based on the important features identified by the optimal model, which is convenient for clinicians and can precisely calculate the risk of DCI after aSAH.

METHODS

Study Design and Patient Enrollment

This multicenter, retrospective, observational cohort study utilized clinical data from the electronic health record system. The study participants consisted of all adult patients with aSAH within 24 h of onset who were treated in the Department of Neurosurgery from April 2019 to June 2021, Renmin Hospital of Wuhan University, Huzhou Central Hospital, Affiliated Hospital of Panzhihua University, General Hospital of Northern Theater Command, First Hospital of Shanxi Medical University. The study eventually enrolled 404 patients (**Figure 1**). According to SAH guidelines, aSAH was diagnosed using head computed tomography (CT), CT angiography, or digital subtraction angiography.

The inclusion criteria were: (1) spontaneous aSAH, (2) admission within 24 h after onset, (3) blood laboratory tests and head CT scans within 24 h after admission, (4) microsurgical clipping or coil embolization within 72 h after onset, and (5) DCI, which occurs within 4–30 days after aSAH.

The exclusion criteria were: (1) admission time exceeded 24 h after onset, (2) intracerebral hemorrhage or vascular

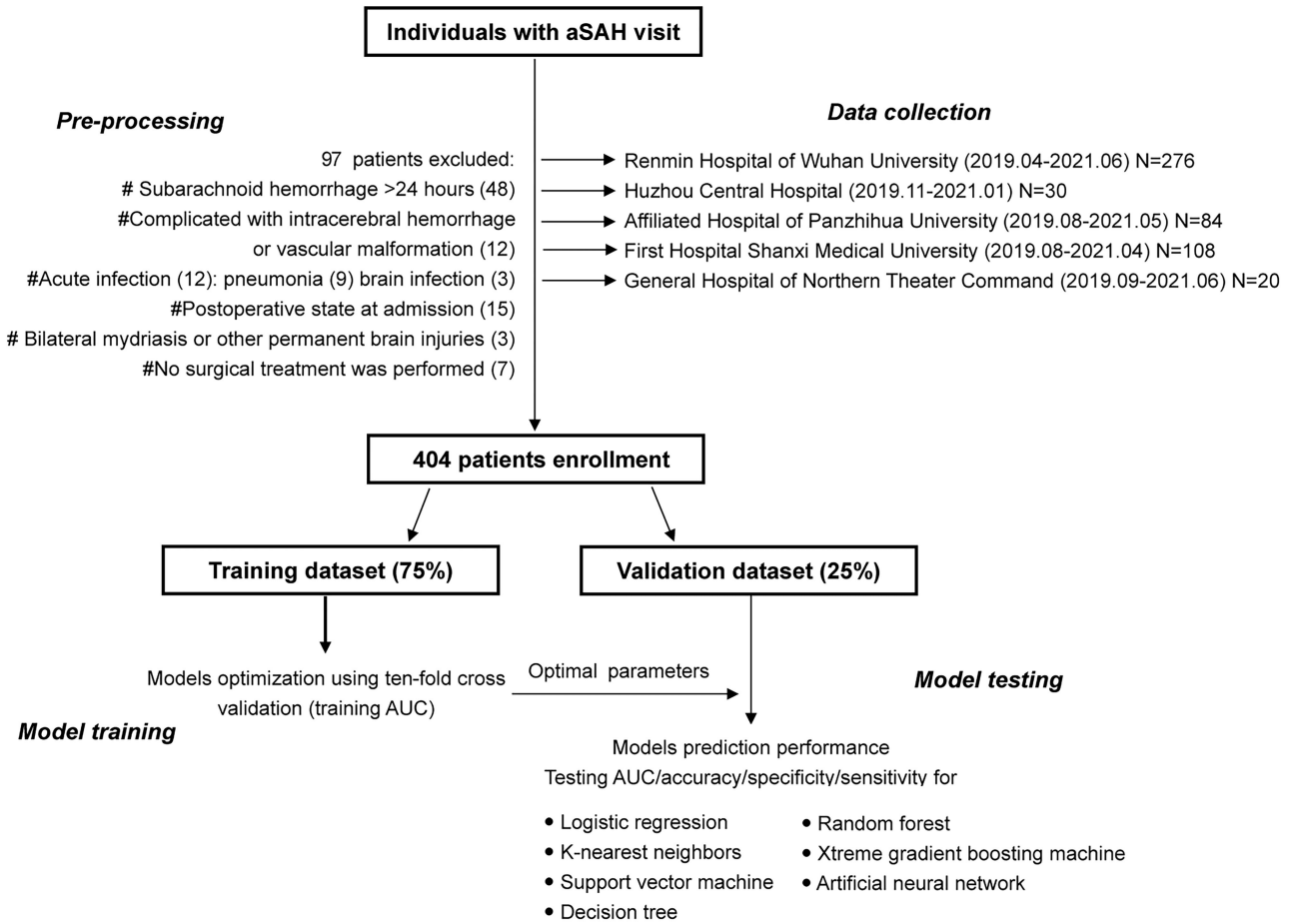


FIGURE 1 | The flowchart of study design and detailed patient enrollment.

malformation, (3) acute infection, (4) postoperative state on admission, (5) bilateral mydriasis or other permanent brain injuries on admission, (6) non-surgical treatment, and (7) patients who died within 3 days after admission.

Clinical Data Collection

Patient demographic data (sex, age), medical history (hypertension, diabetes mellitus, coronary heart disease, smoking, alcohol consumption, anticoagulant treatment, and previous diseases), and clinical state on admission [World Federation of Neurosurgical Societies (WFNS), Hunt and Hess grade (HH), and modified Fisher scale (mFS)] were collected. Aneurysmal details were also recorded, including aneurysm number, location, length, neck size, and treatment. Surgical methods and laboratory tests on admission (glucose, D-dimer, as well as white blood cell [WBC], neutrophil, lymphocyte, and monocyte counts) were also utilized in this study.

CT Value Assessment

The CT values of subarachnoid clots and cerebral edema were manually measured and collected, and measurement methods and references are provided in the **Supplementary Data 1**.

All CT scans were completed using a GE scanner (64-section Optimal CT680) without contrast enhancement. The following parameters were used: tube voltage, 120 kVp; tube current modulation, 300 mA; detector configuration, 64 × 0.625 mm; rotation time, 0.5 s; slice thickness, 5 mm; and collimation, 10 mm.

Regions of interest (ROIs) were manually drawn on the central area of the blood clots in representative slices by two neurosurgeons who were blinded to the patients' clinical information. The mean blood clot density in the subarachnoid space was measured in each ROI (a circle 3–8 mm across), returning the mean Hounsfield Unit (HU) value. Subarachnoid cisterns/fissures, including the lateral Sylvian fissure, anterior interhemispheric fissure, medial Sylvian fissure, suprasellar cistern, ambient cistern, and quadrigeminal cistern, were used to determine the mean HU (Woo et al., 2017; Kanazawa et al., 2020).

Regions of interest (circles 5–10 mm across) of the cerebral edema were bilaterally and symmetrically drawn on a representative CT slice. If blood clots were below the insular cortex, the ROI was drawn on the thalamus and basal ganglia. Otherwise, the ROI was drawn on the bilateral centrum semiovale (Claassen et al., 2002; Ahn et al., 2018).

Outcome Definitions

The definition of DCI should meet at least one of the following criteria: (1) no other etiology could have caused a permanent or temporary focal neurological impairment (such as aphasia, apraxia, hemianopia, or neglect) between 4 and 14 days after aSAH; (2) the Glasgow Coma Scale score decreased by at least two points [either on one of its components (eye opening, verbal response, motor response), or on total score]; and (3) head CT scans revealed a low-density area that was not noticeable on admission or immediately after the operation, and there were no other causes except vasospasms between 4 and 30 days after aSAH (Vergouwen et al., 2010).

Sample Size

We used the events per variable criterion with a value of 10 (Peduzzi et al., 1996) to estimate the effective sample size in this study. Our preliminary analysis indicated that nine variables were entered into a multivariable LR model. Hence, at least 90 patients with DCI should be included in the training cohort. Moreover, according to the risk of DCI occurrence after SAH, ∼30% worldwide, there should be at least 300 patients in the model training cohort.

Processing of Missing Data

This dataset included 17 patients with missing values, which accounted for <5% of the study population, so we directly used the missing value deletion method to process the data (Eekhout et al., 2012).

Model Development

A total of 404 patients with aSAH from five medical centers were prospectively enrolled. We randomly divided the patients as training cohort ($N = 303$) and validation cohort ($N = 101$) according to a ratio of 75–25%. The training cohort was utilized to develop a conventional LR, k-nearest neighbor (KNN), support vector machine, decision tree, random forest (RF), extreme gradient boosting, and artificial neural network (ANN) models.

Machine Learning Models Development

LR

The model was trained by fitting the predictor variables with $P < 0.1$ in univariate analysis to multivariate logistic analysis. We used the backward stepwise regression method based on the Akaike information criterion to select the optimal variables and constructed a final LR model. "MASS" package in R software was performed to fit the model.

LASSO

LASSO regression, which is suitable for analyzing high-dimensional data, was used to select the most informative prediction variables. We used the "glmnet, corrplot, caret" packages and 10-fold cross-validation to obtain the optimal λ and factors.

KNN

KNN model uses local geographic information in the predictive environment to predict the results of the new samples. For example, a KNN model with ten neighbors uses the ten closest observations in multidimensional space to predict the results of a new sample based on a distance assessment. The optimal K value was determined by 10-fold cross-validation and the "e1071, class, kknn, kernlab, caret" packages.

SVM

The uniqueness of SVM algorithms is that they mainly use data points from each result class that is closest to the class boundary or misclassified when determining the boundary structure. The radial basis function was applied in this work, and the optimal gamma value and minimum error of the SVM model were determined by 10-fold cross-validation.

DT

DT algorithms partition the sample data by splitting prediction features at discrete cut-points and are usually presented in the form of a tree. In this study, the decision tree algorithm uses the Gini index to determine each split's optimal variable and location. The cost complexity parameter that penalizes more complex trees is used to control the size of the final tree. Ten-fold cross-validation and "rpart, partykit, caret" packages were used to determine the minimum error value.

RF

RF builds a predictive model by sampling objects and variables, generating multiple decision trees, and classifying objects in turn. Finally, the classification results of each decision tree are summarized, and the mode category in all prediction categories is the category of the object predicted by the RF model. The optimal number of trees was determined using 10-fold cross-validation and "randomForest" package.

XGBoost

XGBoost is an optimized distributed gradient enhancement library designed to be efficient, flexible, and portable. It implements ML algorithms under the Gradient Boosting framework. The optimal parameters were determined by "xgboost" package and 10-fold cross-validation.

ANN

ANN is an algorithmic mathematical model that imitates the behavioral characteristics of animal neural networks and performs distributed and parallel information processing. This kind of network relies on the system's complexity, adjusts the interconnection between a mass of internal nodes to achieve the purpose of processing information, and has the ability of self-learning and self-adaptation. Ten-fold cross-validation and "caret, MASS, neuralnet, vcd" packages were conducted to determine the optimal parameters of this model.

Dynamic Nomogram

A web-based dynamic nomogram application was then developed based on the optimal prediction variables based on optimal features. Calibration curve with 1,000 resample bootstrap was used to assess the calibration ability, and the clinical effectiveness was evaluated by decision curve analysis (DCA) and clinical impact curve (CIC). The packages "rmda,

TABLE 1 | Patinets baseline characteristics in training and validation cohorts.

Characteristics	Training cohort ($n = 303$)	Validation cohort ($n = 101$)	*P*-value
Demographics			
Age (years)	57 (51, 64)	57 (51, 63)	0.903
Gender (Female)	179 (59)	68 (67)	0.175
Medical history			
Hypertension	142 (47)	50 (50)	0.73
Diabetes	9 (3)	3 (3)	1.000
CHD	12 (4)	2 (2)	0.532
Smoking	54 (18)	17 (17)	0.94
Drinking	37 (12)	13 (13)	1.000
Anticoagulant	11 (4)	3 (3)	1.000
Disease history			0.886
ICH	3 (1)	0 (0)	
CI	6 (2)	1 (1)	
WFNS grade			0.163
I–II	227 (75)	68 (67)	
III	37 (12)	13 (13)	
IV	22 (7)	12 (12)	
V	17 (6)	8 (8)	
Hunt and Hess grade			0.327
I–II	211 (70)	59 (58)	
III	59 (19)	27 (27)	
IV	18 (6)	9 (9)	
V	15 (5)	6 (6)	
Modified Fisher scale			0.037
1–2	157 (52)	36 (35)	
3	76 (25)	37 (37)	
4	70 (23)	28 (28)	
Aneurysm location			0.694
ACA	273 (90)	93 (92)	
PCA	30 (10)	8 (8)	
Aneurysm number			0.428
Single	269 (89)	86 (85)	
Multiple (≥2)	34 (11)	15 (15)	
Mean aneurysm size			
Neck (mm)	3.2 (2.4, 3.75)	3.5 (2.5, 4.3)	0.068
Length (mm)	4.4 (3.15, 5.5)	4.9 (3.5, 6.7)	0.017
Aneurysm treatment			0.015
Clipping	154 (51)	66 (65)	
Coiling	149 (49)	35 (35)	
Decompressive craniectomy	16 (5)	11 (11)	0.084
Admission laboratory results			
Glucose (mmol/L)	6.96 (5.94, 8.17)	6.8 (5.61, 8.06)	0.396
D-dimer (mg/L)	1.17 (0.58, 2.5)	1.36 (0.82, 2.5)	0.288
WBC (10^9/L)	11.23 (9.27, 13.99)	11.21 (8.98, 14.4)	0.646
Neutrophil (10^9/L)	9.57 (7.56, 12.28)	9.9 (7.27, 12.8)	0.887
Lymphocyte (10^9/L)	0.9 (0.68, 1.25)	0.92 (0.65, 1.27)	0.842
Monocytes (10^9/L)	0.5 (0.34, 0.7)	0.5 (0.33, 0.7)	0.655
Admission CT value (HU)			
ClotCT	57 (52, 62.02)	58 (54, 63)	0.17
EdemaCT	26.82 (24.2, 28.98)	26.88 (25.07, 29)	0.36
DCI	85 (28)	27 (27)	0.898

DCI, indicates delayed cerebral ischemia; ACA, aneurysm includes anterior cerebral artery, middle cerebral artery, internal cerebral artery, anterior communicating artery; posterior communicating artery; PCA, aneurysm includes posterior cerebral artery, basilar artery, anterior inferior cerebellar artery, posterior inferior cerebellar artery, vertebral artery; ACA, anterior circulation aneurysm; PCA, posterior circulation aneurysm; WBC, White blood cell; CT, computer tomography; HU, Hounsfield Unit; WFNS, World Federation of Neurosurgical Surgeons; ICH, Intracerebral hemorrhage; CI, cerebral infarction; CHD, Coronary heart disease.

MASS, survival, ggplot2, ggridges, DynNom, and riskRegression" and "shinyapps.io" were performed to achieve this process.

Model Performance Evaluation

We used the area under the receiver operating characteristic curve (AUC) with 95% confidence intervals (95% CIs), accuracy, balanced accuracy, confusion matrix, sensitivity, and specificity indicators in both training and validation cohorts to evaluate model performance. The AUC value was used to assess model discrimination, while the calibration curve with 10-fold cross-validation (1,000 resample) and Hosmer–Lemeshow test can reflect the model calibration performance.

Statistical Analysis

We applied the Kolmogorov–Smirnov test to determine the data distribution before formally analyzing the data. Continuous variables analyzed using the independent *t*-test or Mann-Whitney *U*-test are presented as mean ± SD or median with interquartile range. Categorical variables analyzed using the chi-square or Fisher's exact tests are expressed as numbers (percentages). The statistical difference between the AUCs of these models was completed by DeLong test. The feature importance was calculated by Gini index using RF algorithm. The total score of all feature importance was added up to 100. A higher importance coefficient commonly indicated a stronger influence on the occurrence of DCI. For continuous variables that were important for DCI indicator, we used the Youden index to calculate the cut-off value to distinguish patients who were prone to be DCI. All statistical tests were two-tailed and $p < 0.05$ were considered statistically significant. Statistical analyses were conducted using IBM SPSS Statistics for Windows, version 26.0, (IBM Corp., Armonk NY, USA) and R software, version R×64 4.1.0 (https://www.r-project.org/).

RESULTS

Baseline Characteristics

The number of patients with DCI were 85 (28%) and 27 (27%) in training and validation cohorts, and women comprised 179

TABLE 2 | Patients baseline characteristics in model training cohort.

Characteristics	Total ($n = 303$)	Non-DCI ($n = 218$)	DCI ($n = 85$)	*P*-value
Demographics				
Age (years)	57 (51, 64)	57 (52, 65)	56 (49, 64)	0.414
Gender (Female)	179 (59)	126 (58)	53 (62)	0.552
Medical history				
Hypertension	142 (47)	103 (47)	39 (46)	0.932
Diabetes	9 (3)	6 (3)	3 (4)	0.714
CHD	12 (4)	5 (2)	7 (8)	0.042
Smoking	54 (18)	35 (16)	19 (22)	0.263
Drinking	37 (12)	26 (12)	11 (13)	0.962
Anticoagulant	11 (4)	6 (3)	5 (6)	0.19
Disease history				0.367
ICH	3 (1)	2 (1)	1 (1)	
CI	6 (2)	3 (1)	3 (4)	
WFNS grade				<0.001
I–II	227 (75)	185 (85)	42 (49)	
III	37 (12)	22 (10)	15 (18)	
IV	22 (7)	7 (3)	15 (18)	
V	17 (6)	4 (2)	13 (15)	
Hunt and Hess grade				<0.001
I–II	211 (70)	167 (77)	44 (52)	
III	59 (19)	42 (19)	17 (20)	
IV	18 (6)	5 (2)	13 (15)	
V	15 (5)	4 (2)	11 (13)	
Modified Fisher scale				<0.001
1–2	157 (52)	126 (57)	31 (36)	
3	76 (25)	56 (26)	20 (24)	
4	70 (23)	36 (17)	34 (40)	
Aneurysm location				0.695
ACA	273 (90)	195 (89)	78 (92)	
PCA	30 (10)	23 (11)	7 (8)	
Aneurysm number				1.000
Single	269 (89)	194 (89)	75 (88)	
Multiple (≥2)	34 (11)	24 (11)	10 (12)	
Mean aneurysm size				
Neck (mm)	3.2 (2.4, 3.75)	3.2 (2.5, 3.98)	3.2 (2.2, 3.7)	0.228
Length (mm)	4.4 (3.15, 5.5)	4.4 (3.26, 5.5)	4.2 (3, 6)	0.963
Aneurysm treatment				0.002
Clipping	154 (51)	98 (45)	56 (66)	
Coiling	149 (49)	120 (55)	29 (34)	
Decompressive craniectomy	16 (5)	3 (1)	13 (15)	<0.001
Admission laboratory results				
Glucose (mmol/L)	6.96 (5.94, 8.17)	6.94 (5.92, 8.14)	6.96 (6.1, 8.3)	0.454
D-dimer (mg/L)	1.17 (0.58, 2.5)	1.11 (0.55, 2.33)	1.51 (0.78, 3.82)	0.024
WBC (10^9/L)	11.23 (9.27, 13.99)	10.5 (8.75, 12.65)	14.6 (11.7, 17.3)	<0.001
Neutrophil (10^9/L)	9.57 (7.56, 12.28)	8.89 (7.27, 11.16)	12.1 (9.67, 14.7)	<0.001
Lymphocyte (10^9/L)	0.9 (0.68, 1.25)	0.94 (0.68, 1.26)	0.83 (0.68, 1.09)	0.266
Monocytes (10^9/L)	0.5 (0.34, 0.7)	0.44 (0.32, 0.66)	0.65 (0.44, 0.9)	<0.001
Admission CT value (HU)				
ClotCT	57.1 ± 7.05	55.23 ± 6.33	61.9 ± 6.54	<0.001
EdemaCT	26.82 (24.2, 28.98)	26.75 (24.25, 28.45)	27 (24.2, 30)	0.28

ACA, aneurysm includes anterior cerebral artery, middle cerebral artery, internal cerebral artery, anterior communicating artery; posterior communicating artery; PCA, aneurysm includes posterior cerebral artery, basilar artery, anterior inferior cerebellar artery, posterior inferior cerebellar artery, vertebral artery; ACA, anterior circulation aneurysm; PCA, posterior circulation aneurysm; WBC, White blood cell; CT, computer tomography; HU, Hounsfield Unit; WFNS, World Federation of Neurosurgical Surgeons; ICH, Intracerebral hemorrhage; CI, cerebral infarction; CHD, Coronary heart disease.

(59%) and 68 (67%) patients in the two groups, respectively. The median age in both the cohorts was 57 years. In terms of other admission clinical features, there were more patients with mFS of 3–4 point in the validation cohort than training cohort ($P < 0.05$), and the aneurysm mean length size in the validation cohort was larger than the training cohort ($p < 0.05$). Among the patients with aSAH in the validation cohort, there is a larger proportion of patients who chose aneurysm clipping ($p < 0.05$). However, there were no significant differences in medical history, disease history, other clinical conditions, aneurysm location, aneurysm number, admission laboratory results, and admission CT value between the two cohorts ($P > 0.05$). **Table 1** shows the detailed baseline characteristics of the datasets. We also analyzed the baseline characteristics of the DCI and non-DCI groups in the training cohort. **Table 2** shows the detailed baseline data of the two groups in the training cohort.

Model Performance Evaluation and Comparison

When using the validation cohort to evaluate and compare model performance, our results showed that conventional LR with an AUC value of 0.824 (95%CI: 0.73–0.91) outperformed KNN, decision tree, support vector machine, and extreme gradient boosting model with the AUCs of 0.792 (95%CI: 0.68–0.9, DeLong: $P = 0.46$), 0.675 (95%CI: 0.56–0.79, DeLong: $P < 0.01$), 0.677 (95%CI: 0.57–0.77, DeLong: $P < 0.01$), and 0.78 (95%CI: 0.68–0.87, DeLong: $P = 0.50$). However, the RF and ANN model

TABLE 3 | Model performance evaluation using training and validation cohorts.

Cohort	Model	AUC (95%CI)	Accuracy	Sensitivity	Specificity
Training	LR	0.837 (0.784–0.889)	0.828	0.552	0.935
	KNN	0.992 (0.985–0.998)	0.904	0.658	1.000
	SVM	0.781 (0.728–0.834)	0.934	0.765	1.000
	DT	0.827 (0.772–0.883)	0.808	0.623	0.881
	RF	1.000 (1.000–1.000)	1.000	1.000	1.000
	XGB	0.884 (0.844–0.925)	0.828	0.917	0.600
	ANN	1.000 (1.000–1.000)	0.759	0.548	0.905
Validation	LR	0.824 (0.737–0.912)	0.802	0.444	0.932
	KNN	0.792 (0.683–0.901)	0.802	0.407	0.946
	SVM	0.677 (0.578–0.775)	0.772	0.259	0.959
	DT	0.675 (0.559–0.791)	0.703	0.444	0.797
	RF	0.858 (0.783–0.932)	0.802	0.518	0.905
	XGB	0.780 (0.686–0.874)	0.693	0.77	0.481
	ANN	0.858 (0.782–0.932)	0.594	0.365	0.837

LR, logistic regression; KNN, K-nearest neighbor; SVM, support vector machine; DT, decision tree model; RF, random forest; XGBoost, extreme gradient boosting; ANN, artificial neural network.

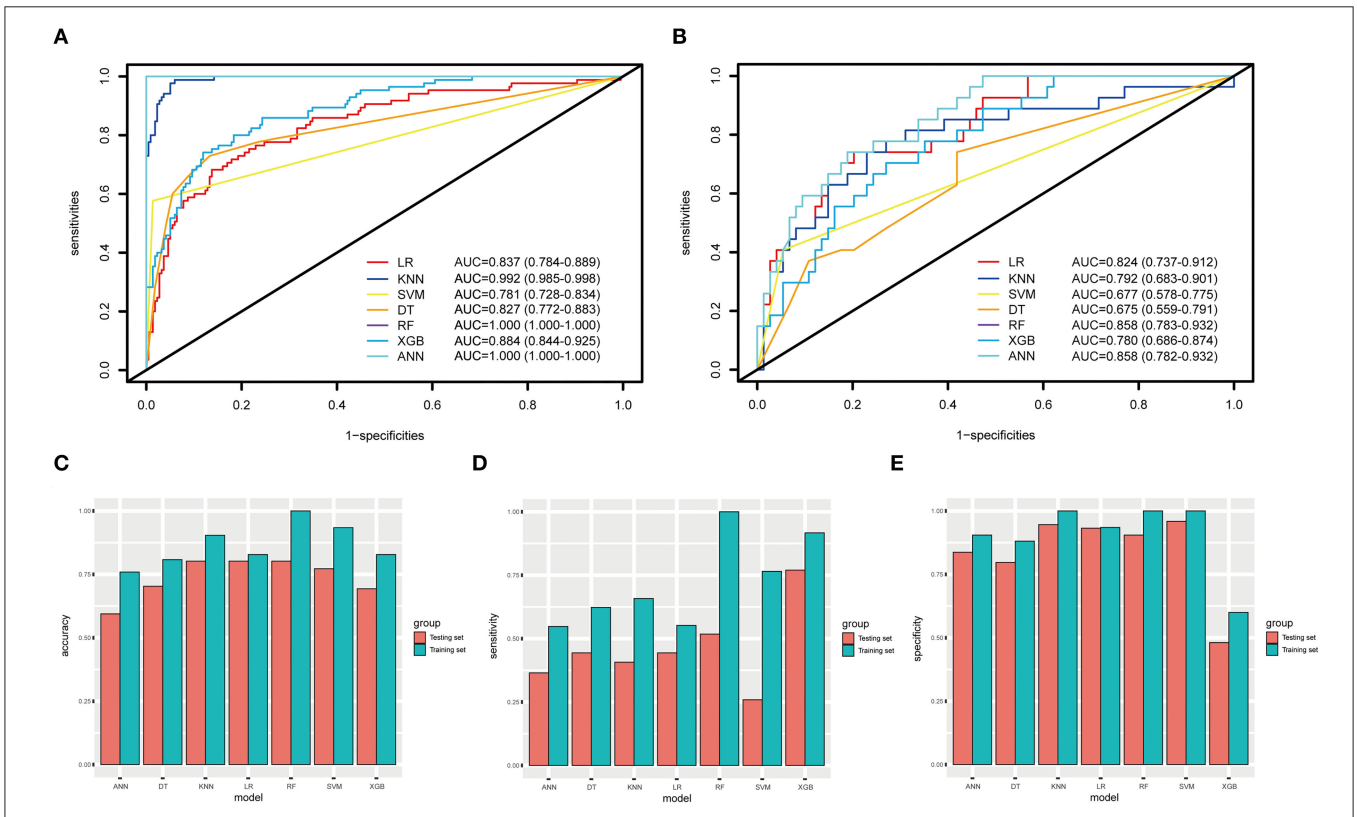


FIGURE 2 | The ROC curves, accuracy, sensitivity, and specificity of ML methods and conventional LR. **(A,B)** The AUCs of ML methods and LR in the training and testing datasets. **(C–E)** The accuracy, sensitivity, and specificity of LR, KNN, SVM, DT, RF, XGB, and ANN in the training and testing datasets are 82.8, 90.4, 93.4%, 80.8, 100, 82.8, 75.9% (blue); 80.2, 80.2, 77.2, 70.3, 80.2, 69.3, 59.4% (orange); 55.2, 65.8, 76.5, 62.3, 100, 91.7, 54.8% (blue); 44.4, 40.7, 25.9, 44.4, 51.8, 77, 36.5% (orange), 93.5, 100, 100, 88.1, 100, 60, 90.5% (blue); 93.2,94.6, 95.9, 79.7, 90.5, 48.1, 83.7% (orange).

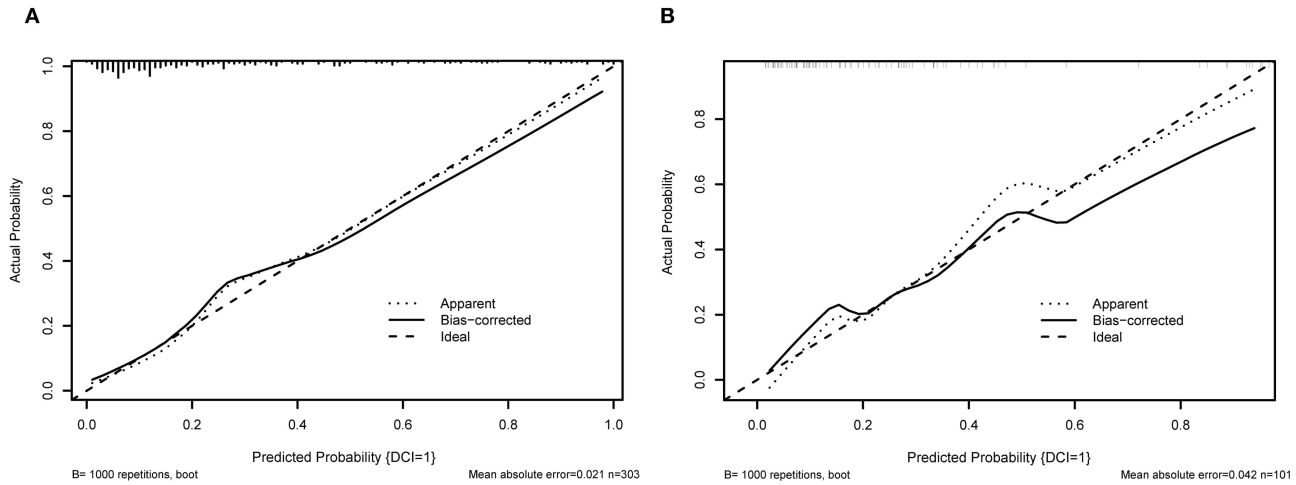


FIGURE 3 | The calibration curve of RF model. **(A)** shows the 10-fold cross validation using training cohort; **(B)** illustrates the 10-fold cross-validation using the validation cohort.

with a same AUC (0.858, 95%CI: 0.78–0.93, DeLong: $P = 0.26$) still performed well than the LR. Furthermore, the accuracy and balanced accuracy of the RF were 20.8 and 11% higher than the latter. **Supplementary Table 1** shows the confusion matrix and balanced accuracy of ML and LR model using training and validation cohorts. **Figure 2** and **Table 3** present the performances of all models when using the training and validation cohorts. In addition, **Figure 3** demonstrates that the superior RF model had a good calibration performance according to the calibration curve and Hosmer–Lemeshow test in the training ($X^2 = 8.78$, df $= 8$, P-value $= 0.36$;) and validation cohort ($X^2 = 10.97$, df $= 8$, P-value $= 0.203$). **Table 4** and **Figure 4** show the process of model development.

Individual Variable Importance

The five most important features for DCI prediction were CT value of subarachnoid hemorrhage (15.68), WBC count (13.72), neutrophil count (12.28), CT value of cerebral edema (8.54), and monocyte count (7.54). The cut-off value of WBC, neutrophil, and monocyte counts for predicting DCI were $11.2 \times 10^{\wedge 9}$/L, $9.58 \times 10^{\wedge 9}$/L, and $0.46 \times 10^{\wedge 9}$/L, respectively. Moreover, the cut-off value of CT value in subarachnoid hemorrhage and cerebral edema were 60.12 (HU) and 28.15 (HU). **Figure 5** shows all input feature importance. An online prediction tool (https://dynamic-nomogram.shinyapps.io/DynNomapp-DCI/) was developed based on the five optimal predictors in the RF model, which could precisely calculate the risk value of DCI after aSAH. A risk percentage of 50% calculated by this tool commonly represents an occurrence of DCI in patients with aSAH. **Figure 6** displays the interface of the online tool for predicting DCI. Both decision curve analysis and clinical impact curve on the validation cohort showed a superior overall net benefit over the entire range of threshold probabilities (**Figure 7**).

TABLE 4 | The univariate and multivariate analysis during fitting logistic regression model.

Variable	OR (95%CI)	*P*-value	Variable	aOR (95%CI)	*P*-value
CHD	3.82 (1.18–13.25)	0.025	*NA*	*NA*	*NA*
WFNS	2.07 (1.65–2.63)	<0.001	WFNS	1.53 (1.11–2.12)	0.009*
HH	1.97 (1.53–2.58)	<0.001	*NA*	*NA*	*NA*
MFS	1.57 (1.24–2.02)	<0.001	MFS	0.76 (0.53–1.08)	0.137
Treatment	0.42 (0.24–0.71)	0.001	Treatment	0.41 (0.21–0.77)	0.007*
DC	12.93 (4.03–57.6)	<0.001	DC	4.54 (1.01–25.13)	0.059
ClotCT	1.17 (1.12–1.23)	<0.001	ClotCT	1.11 (1.05–1.17)	<0.001*
WBC	1.29 (1.19–1.39)	<0.001	WBC	1.58 (1.11–2.33)	0.018*
NC	1.25 (1.16–1.36)	<0.001	NC	0.74 (0.5–1.07)	0.141
MC	8.44 (3.48–21.59)	<0.001	*NA*	*NA*	*NA*
D-dimer	1.07 (1.02–1.15)	0.016	*NA*	*NA*	*NA*

*CHD, Coronary heart disease; WFNS, World Federation of Neurosurgical Surgeons; HH, Hunt and Hess grade; MFS, modified Fisher scale; DC, Decompressive craniectomy; WBC, White blood cell; NC, Neutrophil; MC, Monocytes. aOR, adjusted odds ratio. *indicates statistical significance ($p < 0.05$) by multivariate logistic regression.*

DISCUSSION

In this study, the eligible patients with aSAH from five medical centers were randomly divided into model training and validation cohorts. One conventional LR and six types of famous ML methods were used to construct the prediction model by incorporating relatively complete admission clinical data, and all model performances were assessed and compared. To the best of our knowledge, this study is the first to utilize the rounded clinical features to develop the model and systematically compare the performance of several popular ML methods and conventional LR on DCI prediction. In addition, firstly, we developed an online prediction tool based on the most important features of the RF model to precisely calculate the risk of DCI development.

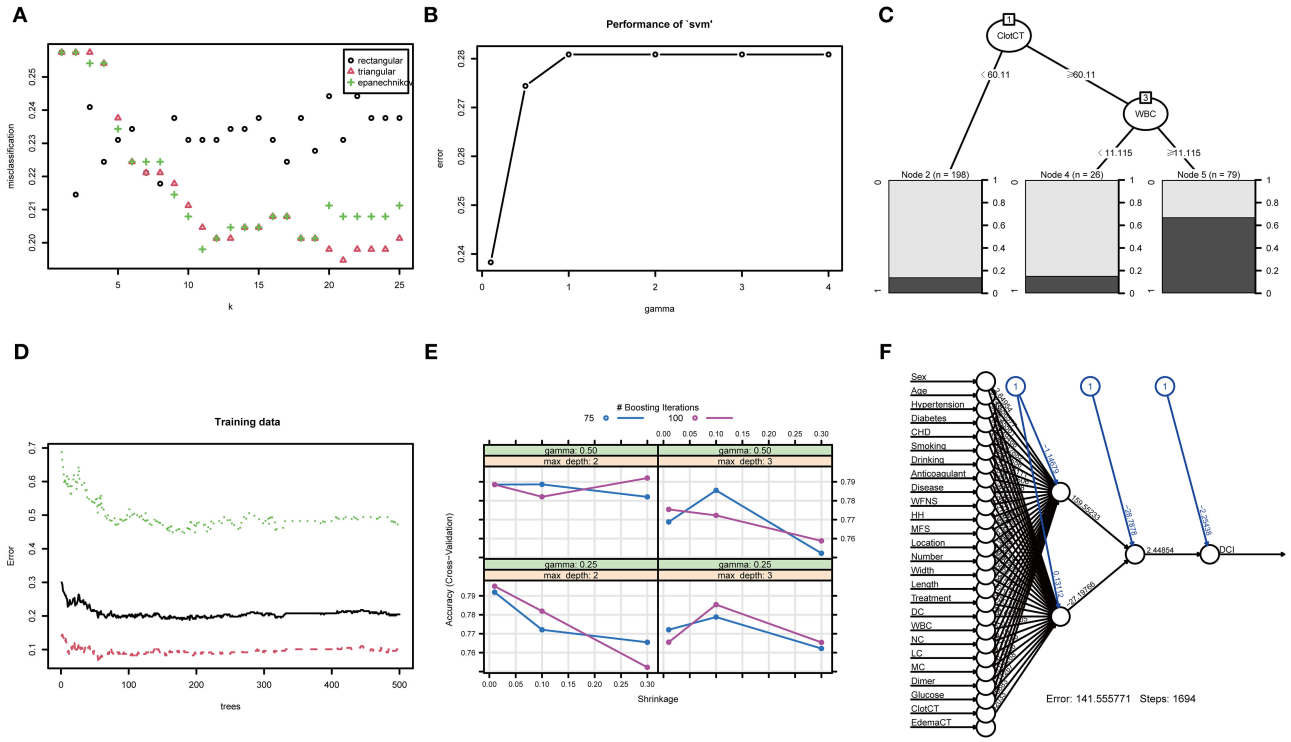


FIGURE 4 | The training process and optimal parameters of six types of full-feature ML models. **(A)** shows the training process of k-nearest neighbor, and the optimal *K*-value is 21; **(B)** shows the error and gamma parameter of the support vector machine during the training process, and the two optimal parameters are 0.238 and 0.1; **(C)** illustrates the training process of the decision tree model, and the most important branches are subarachnoid clot CT value and WBC count; **(D)** displays the training process of random forest, and the optima tree number of the RF model is 179; **(E)** shows the training process of eXtreme Gradient Boosting, and the optimal parameters are gamma of 0.25, max depth of 2, and n-rounds value of 100; **(F)** demonstrates the training process of artificial neural network, and generalized weights of all clinical features are seen.

It was considered that only a few admission clinical features would not lead to an accurate DCI prediction. However, the most commonly used multivariable prediction models are still based on LR. For instance, de Rooij et al. (2013) incorporated some features selected by LR and constructed a practical risk chart for DCI prediction. The AUC value of this risk chart was 0.69 in the validation cohort. Liu et al. (2020) used six factors selected *via* LR to develop a nomogram for DCI, which achieved an AUC value of 0.65 on the test set. Other studies have also employed the conventional LR method to identify independent factors for DCI prediction (Al-Mufti et al., 2017, 2019a; Duan et al., 2018; Hurth et al., 2020). In our study, the LR model incorporated four independent features for DCI classification and achieved an AUC value of 0.837 in the validation cohort, which was higher than the AUC values previous models reported (de Oliveira Manoel et al., 2015; Foreman et al., 2017; van der Steen et al., 2019; Liu et al., 2020). The inclusion of complete admission clinical information can enable the LR to select the optimal variables to improve the prediction performance, which may explain the better performance of our LR model. However, owing to the robustness of the LR model, it cannot take full advantage of information from all clinical input features.

Machine learning models can solve the problem of high-dimensional data more robustly than the conventional LR method, making them suitable for fitting more features for prediction (Brusko et al., 2018; Buchlak et al., 2020). This capability can reduce the subjectivity in statistical analysis and ensure the objectivity of the results. Recently, ML algorithms have been developed rapidly, and some studies have reported the use of ML to predict the occurrence of DCI. de Jong et al. (2021) constructed a feedforward artificial neural network model and achieved an AUC of 0.72 for DCI prediction with a database with 362 patients. Their model performed equally well as the VASOGRADE model (de Oliveira Manoel et al., 2015). The ANN model in our study, with an AUC of 0.858, had a better predictive power than the conventional LR model and outperformed the previous ANN model.

Some researchers have compared the performances of LR and ML models for the prediction of DCI or other diseases. For instance, Savarraj et al. (2021) developed ML and LR models for DCI classification using a dataset with 399 patients. Their results showed that the ML model with the highest AUC value of 0.75 ± 0.07 outperformed the LR model. Ramos et al. (2019) reported that the ML model with the highest AUC value of 0.74 performed better than the best LR model with an AUC of 0.63. However, Nusinovici et al. (2020) reported that the LR model could perform equivalently to the ML models in their study, and Chen et al. (2020) showed that ML models

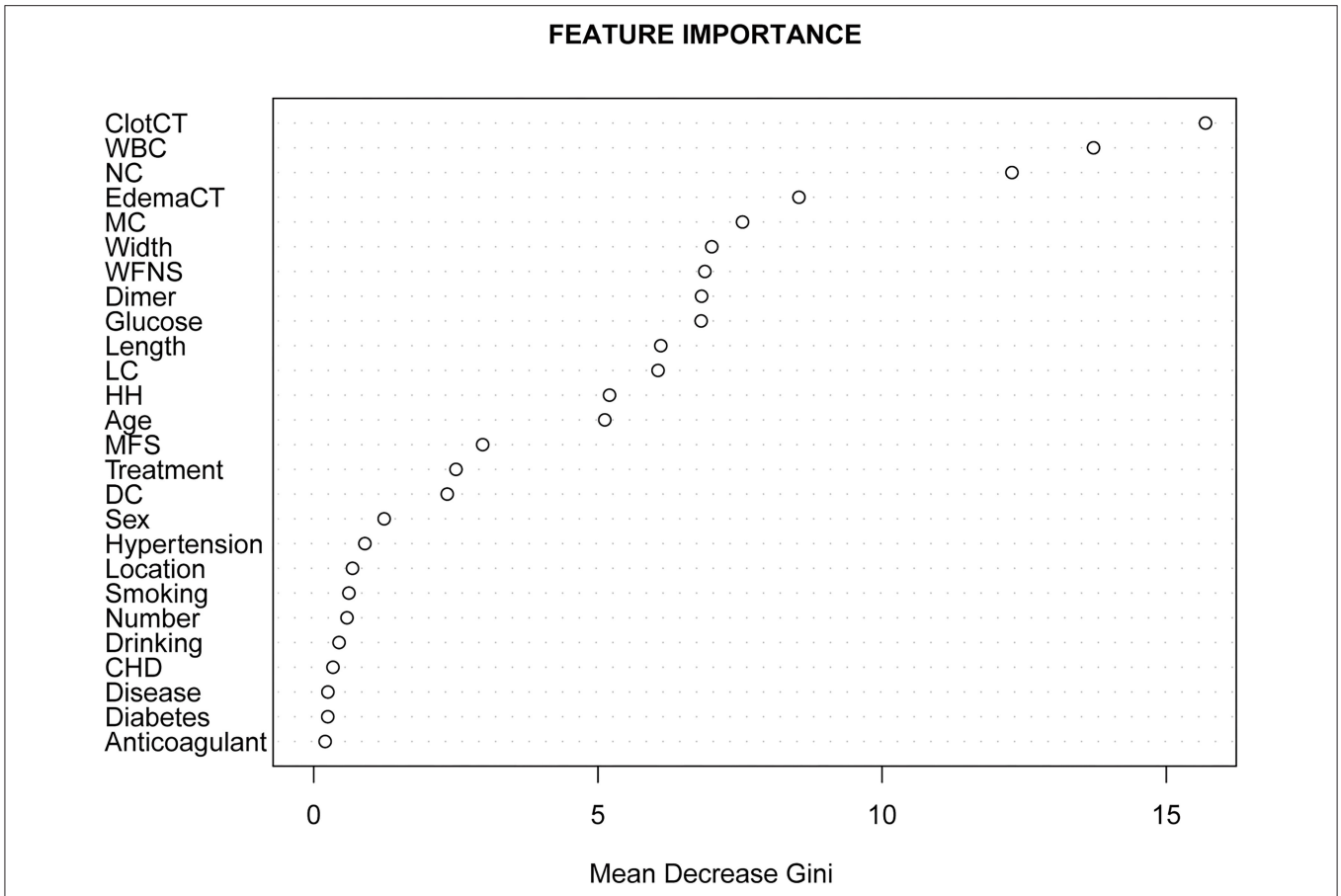


FIGURE 5 | All features importance of random forest calculated by Gini index.

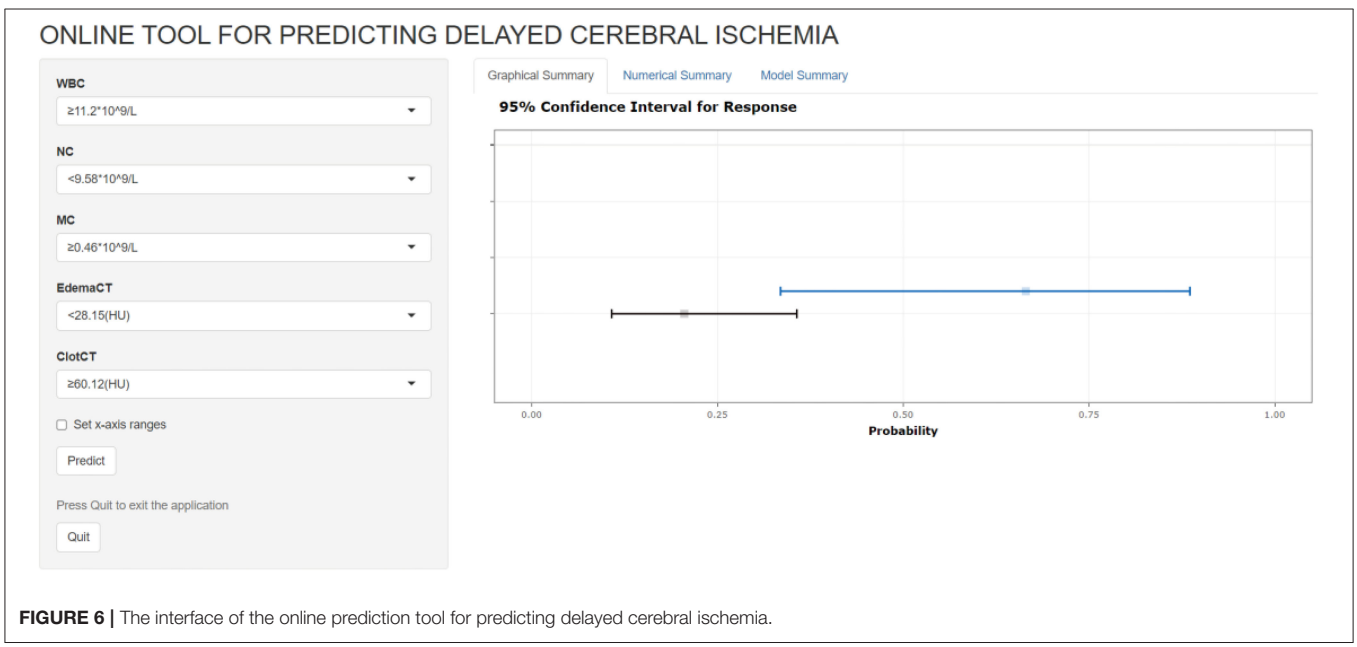


FIGURE 6 | The interface of the online prediction tool for predicting delayed cerebral ischemia.

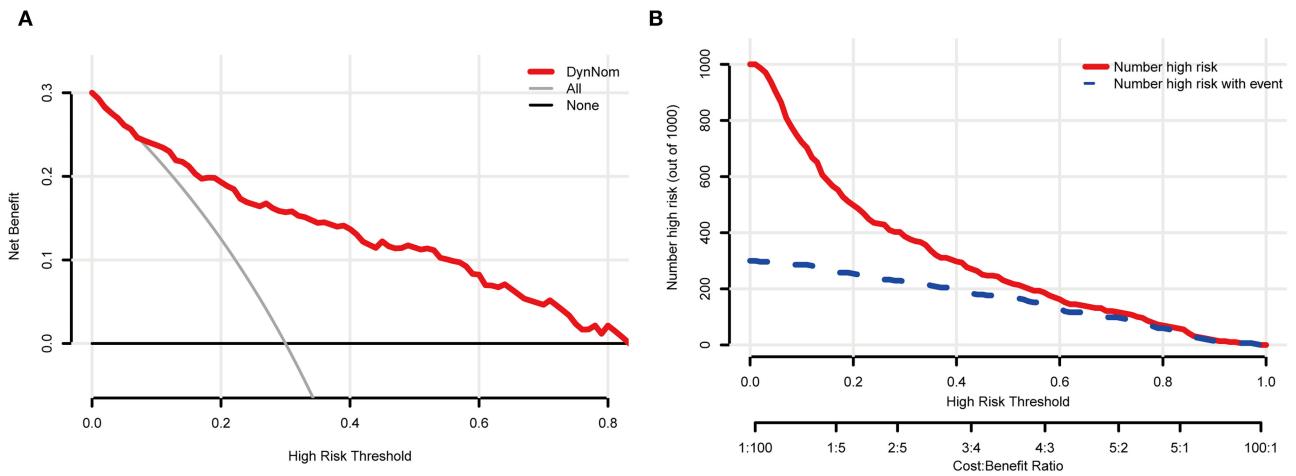


FIGURE 7 | The clinical and practical evaluation of the online prediction tool. **(A)** shows a decision curve analysis (DCA); **(B)** displays a clinical impact curve (CIC).

cannot outperform the conventional LR model in predicting other diseases. In our study, we constructed several popular ML models based on the relatively complete clinical features, some of which were not compared in previous studies. The prediction ability of the LR model was inferior to those of the ANN and RF models, but better than those of the KNN, support vector machine, decision tree, and extreme gradient boosting models. This indicates that the traditional LR method still can play an important role in DCI prediction. Although ML can make perfect use of the input characteristics, data overfitting may lead to poor prediction performance.

Subarachnoid hemorrhage is a state of systemic inflammatory response syndrome, with both biochemical and cellular reactions (Parkinson and Stephensen, 1984). SAH initiates the rapid activation of the inflammatory cascade, and growing evidence suggests that an early neurovascular inflammatory response is a potential mechanism of late cerebral vasospasm and early brain injury (Helbok et al., 2015). The CT value in SAH often represents the subarachnoid clot density and can reflect the cerebral inflammatory response. At present, the measurement method of CT density value of subarachnoid clot still relies on the manual drawing of ROI. Kanazawa et al. (2020) found that an ROI CT value of $\geq$49.95 HU is correlated with DCI occurrence. Our results are consistent with those of previous studies showing that the CT value of >60.12 HU plays a prominent role in DCI prediction. Additionally, Ahn et al. (2018) constructed a scoring system for predicting DCI and clinical outcomes based on early cerebral edema after aSAH. This scoring system may become a surrogate marker of early brain injury and predicts DCI and prognosis after aSAH. Our consequence also illustrates that early cerebral edema also has an important influence on DCI prediction. As we know, WBC and neutrophil counts also play an important role in reflecting neuroinflammatory responses. Al-Mufti et al. (2019b) found that a WBC count >12.1 $\times$ 10^9/L was the strongest predictor of DCI after adjusting for confounding factors, including clinical grade and aneurysm clipping treatment. Our results found that the WBC count >11.2 $\times$ $10^{\wedge 9}$/L, neutrophil count >9.58 $\times$ $10^{\wedge 9}$/L, and monocytes count >0.46 $\times$ $10^{\wedge 9}$/L were the most important features for the prediction DCI. A recent study has shown that admission WBC, neutrophil, and monocyte counts were higher in patients with DCI and unfavorable prognosis (Gusdon et al., 2021). Inspiringly, our study confirmed this, which could account for the fact that DCI development is closely relative to the inflammatory response. Future basic research should further explore the inflammatory machine during the occurrence of DCI.

Based on the superior prediction performance of the RF, we used the most important features to construct an online prediction tool, which will aid in the early identification of patients at high risk of DCI after aSAH and allow timely interventions.

Our study systematically collected admission baseline information, laboratory test results, and admission CT imaging data, and these pieces of information are representative as possible of the true condition of aSAH patients when they are admitted to the hospital. Secondly, in order to avoid the defects of single-center data modeling, we collected data from multiple medical centers, making the DCI prediction model more generalized and robust, which is the second innovation of this study. Thirdly, this study covers several of the most popular machine learning algorithms, which have not been systematically compared with conventional models in previous studies, which is also an innovation point. Fourth, we built an online version of the prediction tool, which is convenient for clinicians to calculate the risk of DCI based on patient information at admission. However, there are several limitations that were observed. This was a retrospective study, and a larger prospective study should be considered to validate our results. Second, a possible deviation caused by manual ROI drawing is unavoidable. The agreement measurements for CT values between an experienced neurosurgeon and a radiologist were acceptable. Third, having

an accuracy of 1 or AUC of 1 on the training dataset means the model is perfect, which is clearly not the case. Among the model we constructed, the random forest has overfitting. We know that overfitting may occur when the model tries to fit all the predicted features with a limited training dataset, which is to say a modeling error in statistics that occurs when a function is too closely aligned to the training dataset. Our future studies will collect more samples to further verify the results of the RF mode.

CONCLUSIONS

In this multicenter study, we found that several ML methods, particularly random forest, outperformed conventional LR. Furthermore, an online prediction tool based on the random forest model was developed to identify patients at high risk for delayed cerebral ischemia after subarachnoid hemorrhage and facilitate timely interventions.

DATA AVAILABILITY STATEMENT

The raw data that supporting the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

The Medical Ethics Committee of Renmin Hospital of Wuhan University (our principal affiliation site) approved the study protocol (approval number WDRM2021-K022). The Ethics Committees of Huzhou Central Hospital (202108005-01), the Affiliated Hospital of Panzhihua University (202105002), General Hospital of Northern Theater Command (Y2021060), and the First Hospital of Shanxi Medical University (2021-Y6) also approved the protocol. The Medical Ethics Committee waived the need for patient consent because the data were derived from the electronic health record system. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

PH, YLi, HZ, and QC: study design. PH, YLi, LY, and YX: literature search. PH, YLi, YLiu, GG, XG, ZS, YS, and QS: data acquisition. PH, SY, YQ, and ML: data analysis and statistical analysis. PH, YLi, HZ, YQ, LW, GD, and QC: manuscript preparation, editing, and review. GD, ML, and XN: funding supporting. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 82001311), the National Natural Science Foundation of China (No. 81971870), the National Natural Science Foundation of China (No. 81703752), and the National Natural Science Foundation of China (No. 82001385).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.857521/full#supplementary-material

REFERENCES

Ahn, S. H., Savarraj, J. P., Pervez, M., Jones, W., Park, J., Jeon, S. B., et al. (2018). The subarachnoid hemorrhage early brain edema score predicts delayed cerebral ischemia and clinical outcomes. *Neurosurgery* 83, 137–145. doi: 10.1093/neuros/nyx364

Al-Mufti, F., Amuluru, K., Damodara, N., Dodson, V., Roh, D., Agarwal, S., et al. (2019a). Admission neutrophil-lymphocyte ratio predicts delayed cerebral ischemia following aneurysmal subarachnoid hemorrhage. *J. Neurointerv. Surg.* 11, 1135–1140. doi: 10.1136/neurintsurg-2019-014759

Al-Mufti, F., Amuluru, K., Smith, B., Damodara, N., El-Ghanem, M., Singh, I. P., et al. (2017). Emerging markers of early brain injury and delayed cerebral ischemia in aneurysmal subarachnoid hemorrhage. *World Neurosurg.* 107, 148–159. doi: 10.1016/j.wneu.2017.07.114

Al-Mufti, F., Misiolek, K. A., Roh, D., Alawi, A., Bauerschmidt, A., Park, S., et al. (2019b). White blood cell count improves prediction of delayed cerebral ischemia following aneurysmal subarachnoid hemorrhage. *Neurosurgery* 84, 397–403. doi: 10.1093/neuros/nyy045

Brusko, G. D., Kolcun, J. P. G., and Wang, M. Y. (2018). Machine-learning models: the future of predictive analytics in neurosurgery. *Neurosurgery* 83, E3–e4. doi: 10.1093/neuros/nyy166

Buchlak, Q. D., Esmaili, N., Leveque, J. C., Farrokhi, F., Bennett, C., Piccardi, M., et al. (2020). Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review. *Neurosurg. Rev.* 43, 1235–1253. doi: 10.1007/s10143-019-01163-8

Chen, G., Lu, M., Shi, Z., Xia, S., Ren, Y., Liu, Z., et al. (2020). Development and validation of machine learning prediction model based on computed tomography angiography-derived hemodynamics for rupture status of intracranial aneurysms: a Chinese multicenter study. *Eur. Radiol.* 30, 5170–5182. doi: 10.1007/s00330-020-06886-7

Claassen, J., Carhuapoma, J. R., Kreiter, K. T., Du, E. Y., Connolly, E. S., and Mayer, S. A. (2002). Global cerebral edema after subarachnoid hemorrhage: frequency, predictors, and impact on outcome. *Stroke* 33, 1225–1232. doi: 10.1161/01.STR.0000015624.29071.1F

de Jong, G., Aquarius, R., Sanaan, B., Bartels, R., Grotenhuis, J. A., Henssen, D., et al. (2021). Prediction models in aneurysmal subarachnoid hemorrhage: forecasting clinical outcome with artificial intelligence. *Neurosurgery* 88, E427–e34. doi: 10.1093/neuros/nyaa581

de Oliveira Manoel, A. L., Jaja, B. N., Germans, M. R., Yan, H., Qian, W., Kouzmina, E., et al. (2015). The VASOGRADE: a simple grading scale for prediction of delayed cerebral ischemia after subarachnoid hemorrhage. *Stroke* 46, 1826–1831. doi: 10.1161/STROKEAHA.115.008728

de Rooij, N. K., Greving, J. P., Rinkel, G. J., and Frijns, C. J. (2013). Early prediction of delayed cerebral ischemia after subarachnoid hemorrhage: development and validation of a practical risk chart. *Stroke* 44, 1288–1294. doi: 10.1161/STROKEAHA.113.001125

Duan, W., Pan, Y., Wang, C., Wang, Y., Zhao, X., Wang, Y., et al. (2018). Risk factors and clinical impact of delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage: analysis from the China National Stroke Registry. *Neuroepidemiology* 50, 128–136. doi: 10.1159/000487325

Eekhout, I., de Boer, R. M., Twisk, J. W., de Vet, H. C., and Heymans, M. W. (2012). Missing data: a systematic review of how they are reported and handled. *Epidemiology* 23, 729–732. doi: 10.1097/EDE.0b013e3182576cdb

Foreman, P. M., Chua, M. H., Harrigan, M. R., Fisher, W. S. 3rd, Tubbs, R. S., Shoja, M. M., et al. (2017). External validation of the Practical Risk Chart for the prediction of delayed cerebral ischemia following aneurysmal subarachnoid hemorrhage. *J. Neurosurg.* 126, 1530–1536. doi: 10.3171/2016.1.JNS 152554

Francoeur, C. L., and Mayer, S. A. (2016). Management of delayed cerebral ischemia after subarachnoid hemorrhage. *Crit. Care* 20, 277. doi: 10.1186/s13054-016-1447-6

Fugate, J. E., and Rabinstein, A. A. (2012). Intensive care unit management of aneurysmal subarachnoid hemorrhage. *Curr. Neurol. Neurosci. Rep.* 12, 1–9. doi: 10.1007/s11910-011-0230-y

Gusdon, A. M., Savarraj, J. P. J., Shihabeddin, E., Paz, A., Assing, A., Ko, S. B., et al. (2021). Time course of peripheral leukocytosis and clinical outcomes after aneurysmal subarachnoid hemorrhage. *Front. Neurol.* 12, 694996. doi: 10.3389/fneur.2021.694996

Helbok, R., Schiefecker, A. J., Beer, R., Dietmann, A., Antunes, A. P., Sohm, F., et al. (2015). Early brain injury after aneurysmal subarachnoid hemorrhage: a multimodal neuromonitoring study. *Crit. Care* 19, 75. doi: 10.1186/s13054-015-0809-9

Hurth, H., Birkenhauer, U., Steiner, J., Schlak, D., Hennersdorf, F., and Ebner, F. H. (2020). Delayed cerebral ischemia in patients with aneurysmal subarachnoid hemorrhage - serum D-dimer and C-reactive protein as early markers. *J. Stroke Cerebrovasc. Dis.* 29, 104558. doi: 10.1016/j.jstrokecerebrovasdis.2019.104558

Jordan, M. I., and Mitchell, T. M. (2015). Machine learning: trends, perspectives, and prospects. *Science* 349, 255–260. doi: 10.1126/science.aaa8415

Kanazawa, T., Takahashi, S., Minami, Y., Jinzaki, M., Toda, M., and Yoshida, K. (2020). Early prediction of clinical outcomes in patients with aneurysmal subarachnoid hemorrhage using computed tomography texture analysis. *J. Clin. Neurosci.* 71, 144–149. doi: 10.1016/j.jocn.2019.08.098

Liu, H., Xu, Q., and Li, A. (2020). Nomogram for predicting delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage in the Chinese population. *J. Stroke Cerebrovasc. Dis.* 29, 105005. doi: 10.1016/j.jstrokecerebrovasdis.2020.105005

Macdonald, R. L. (2014). Delayed neurological deterioration after subarachnoid haemorrhage. *Nat. Rev. Neurol.* 10, 44–58. doi: 10.1038/nrneurol.2013.246

Molyneux, A. J., Kerr, R. S., Yu, L. M., Clarke, M., Sneade, M., Yarnold, J. A., et al. (2005). International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised comparison of effects on survival, dependency, seizures, rebleeding, subgroups, and aneurysm occlusion. *Lancet* 366, 809–817. doi: 10.1016/S0140-6736(05)67214-5

Nusinovici, S., Tham, Y. C., Chak Yan, M. Y., Wei Ting, D. S., Li, J., Sabanayagam, C., et al. (2020). Logistic regression was as good as machine learning for predicting major chronic diseases. *J. Clin. Epidemiol.* 122, 56–69. doi: 10.1016/j.jclinepi.2020.03.002

Parkinson, D., and Stephensen, S. (1984). Leukocytosis and subarachnoid hemorrhage. *Surg. Neurol.* 21, 132–134. doi: 10.1016/0090-3019(84)90330-6

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., and Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* 49, 1373–1379. doi: 10.1016/S0895-4356(96)00236-3

Ramos, L. A., van der Steen, W. E., Sales Barros, R., Majoie, C., van den Berg, R., Verbaan, D., et al. (2019). Machine learning improves prediction of delayed cerebral ischemia in patients with subarachnoid hemorrhage. *J. Neurointerv. Surg.* 11, 497–502. doi: 10.1136/neurintsurg-2018-014258

Savarraj, J. P. J., Hergenroeder, G. W., Zhu, L., Chang, T., Park, S., Megjhani, M., et al. (2021). Machine learning to predict delayed cerebral ischemia and outcomes in subarachnoid hemorrhage. *Neurology* 96, e553–e62. doi: 10.1212/WNL.0000000000011211

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J. Clin. Epidemiol.* 49, 1225–1231. doi: 10.1016/S0895-4356(96)00002-9

van der Steen, W. E., Marquering, H. A., Boers, A. M. M., Ramos, L. A., van den Berg, R., Vergouwen, M. D. I., et al. (2019). Predicting delayed cerebral ischemia with quantified aneurysmal subarachnoid blood volume. *World Neurosurg.* 130:e613–e19. doi: 10.1016/j.wneu.2019.06.170

Vergouwen, M. D., Vermeulen, M., van Gijn, J., Rinkel, G. J., Wijdicks, E. F., Muizelaar, J. P., et al. (2010). Definition of delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage as an outcome event in clinical trials and observational studies: proposal of a multidisciplinary research group. *Stroke* 41, 2391–2395. doi: 10.1161/STROKEAHA.110.589275

Woo, P. Y. M., Tse, T. P. K., Chan, R. S. K., Leung, L. N. Y., Liu, S. K. K., Leung, A. Y. T., et al. (2017). Computed tomography interobserver agreement in the assessment of aneurysmal subarachnoid hemorrhage and predictors for clinical outcome. *J. Neurointerv. Surg.* 9, 1118–1124. doi: 10.1136/neurintsurg-2016-012576

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

PERSPECTIVE
published: 08 July 2022
doi: 10.3389/fnagi.2022.909029

Cognition Meets Gait: Where and How Mind and Body Weave Each Other in a Computational Psychometrics Approach in Aging

*Francesca Bruni[1], Francesca Borghesi[2], Valentina Mancuso[1], Giuseppe Riva[2,3], Marco Stramba-Badiale[4], Elisa Pedroli[1,2] and Pietro Cipresso[2,5]**

[1] Faculty of Psychology, eCampus University, Novedrate, Italy, [2] Applied Technology for Neuropsychology Lab, IRCCS Istituto Auxologico Italiano, Milan, Italy, [3] Human Technology Lab, Catholic University of the Sacred Heart, Milan, Italy, [4] Department of Geriatrics and Cardiovascular Medicine, IRCCS Istituto Auxologico Italiano, Milan, Italy, [5] Department of Psychology, University of Turin, Turin, Italy

Aging may be associated with conditions characterized by motor and cognitive alterations, which could have a detrimental impact on daily life. Although motors and cognitive aspects have always been treated as separate entities, recent literature highlights their relationship, stressing a strong association between locomotion and executive functions. Thus, designing interventions targeting the risks deriving from both components' impairments is crucial: the dual-task represents a starting point. Although its role in targeting and decreasing difficulties in aging is well known, most interventions are focused on a single domain, proposing a vertical model in which patients emerge only for a single aspect per time during assessment and rehabilitation. In this perspective, we propose a view of the individual as a whole between mind and body, suggesting a multicomponent and multidomain approach that could integrate different domains at the same time retracing lifelike situations. Virtual Reality, thanks to the possibility to develop daily environments with engaging challenges for patients, as well as to manage different devices to collect multiple data, provides the optimal scenario in which the integration could occur. Artificial Intelligence, otherwise, offers the best methodologies to integrate a great amount of various data to create a predictive model and identify appropriate and individualized interventions. Based on these assumptions the present perspective aims to propose the development of a new approach to an integrated, multimethod, multidimensional training in order to enhance cognition and physical aspects based on behavioral data, incorporating consolidated technologies in an innovative approach to neurology.

Keywords: virtual reality, machine learning, aging, artificial intelligence, psychometric, rehabilitation, dual-task, neurology

OPEN ACCESS

Edited by:
Wai Haung Yu,
University of Toronto, Canada

Reviewed by:
Laura Mandolesi,
University of Naples Federico II, Italy

****Correspondence:***
Pietro Cipresso
p.cipresso@auxologico.it

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *31 March 2022*
Accepted: *16 June 2022*
Published: *08 July 2022*

Citation:
Bruni F, Borghesi F, Mancuso V, Riva G, Stramba-Badiale M, Pedroli E and Cipresso P (2022) Cognition Meets Gait: Where and How Mind and Body Weave Each Other in a Computational Psychometrics Approach in Aging. Front. Aging Neurosci. 14:909029. doi: 10.3389/fnagi.2022.909029

INTRODUCTION

Population aging is one of the most significant difficulties that social and health systems are facing today. According to the World Health Organization, in 2020 the population older than 60 years old represented 13.5% of the world's population. This number is 2.5 times higher than it was in 1980, and it is growing at an alarming rate (World Health Organization, 2020). The

current situation emphasizes the significance of reacting to the constant rise in aging demand by proposing new solutions to challenges arising from the physiological condition to better meet their needs. A significant proportion of older people may develop frailty, multi-morbidity, and disability causing a significant impact on daily life (Lutz et al., 2008; World Health Organization, 2020). The two most frequent conditions of vulnerability in aging are frailty and cognitive impairment: they are described as a pattern of decline in different aspects of motor (i.e., gait, mobility, balance) and cognitive functions (i.e., attention and working memory) (Gobbens et al., 2010), that could increase disability, mortality, reduce the quality of life, and contribute to adverse outcomes (Speechley and Tinetti, 1991; Fried et al., 2004; Rockwood, 2005; Panza et al., 2015; Kojima et al., 2018; de Morais Fabrício et al., 2020) with a direct impact on health as well as on health care and social costs (World Health Organization, 2020). As a result, the strong links between aging and altered motor and cognitive processes should be highlighted, considering also the well-known link between motor and cognitive functions. An example of this relationship is locomotion. Although locomotion appears to be a mechanical effort, it cannot be reduced to a simple action involving a sequence of repetitive movements. This action requires the coordination of gait and one or more cognitive processes at the same time (Yogev et al., 2008; Al-Yahya et al., 2011). In everyday activities, it may be essential to adjust our gait to overcome environmental impediments or just participate in a discussion; these are all requests that necessitate the ability to accomplish many tasks at the same time. According to the Cognitive-Motor Interference theory (CMI) (Montero-Odasso and Speechley, 2018), performing tasks simultaneously, the so-called dual-task (DT), requires a high level of cognitive control in terms of executive processes and attentional abilities whose impairment may produce a deterioration in either motor and cognitive execution or even both (Yogev et al., 2008). As a result, it appears clear how executive functions are an important cognitive resource for normal locomotion. Executive functions are integrative processes involving cognitive and behavioral components that are required for targeted and successful actions, and attentional resource control which are the foundation of the ability to manage autonomous daily activities (Yogev et al., 2008).

Based on these premises, the need to consider both cognition and gait as intertwined aspects is crucial, as said by Giovenale *mens sana in corpore sano*, considering the individual as a whole between mind and body. Despite this consideration has been well known for centuries and the integration of different domains being daily required, in clinical practice motor and cognitive processes often continue to be treated as separate entities, and the intervention methodologies available are frequently provided vertically, focusing on a single aspect at a time. This is a disadvantage and could be attributed to the limitations in considering deficits in aging, which are confined to be assessed and consequently treated separately concerning motor and cognition as in the case of frailty, that most authors operationalized by focusing predominantly on the physical elements (de Vries et al., 2011). Nowadays, researchers are trying to solve the question relative how a healthy body promotes a healthy mind, hence how a healthful mind could influence a good physical state. In this field, a possible solution appears the so-called DT paradigm (Dorfman et al., 2014; Lauenroth et al., 2016). The DT paradigm involves two exercises performed at the same time, particularly a motor task and a cognitive one (e.g., walking while counting backward). Evidence suggests that DT is a powerful tool during both assessment and rehabilitation processes. On one hand, it is more sensitive than a single task to detect some early gait dysfunctions (Lunardini et al., 2021). On the other hand, it benefits a variety of individuals, including older persons with frailty syndrome, neurological disease, and poststroke patients (Deutsch et al., 2013; Wajda et al., 2017; Freitag et al., 2019). However, the most available studies focus on its benefits only for a healthy body—namely, its implications on physical processes. For instance, substantial differences have been shown when comparing CMI with single-task exercise or no intervention on physical outcomes (Wang et al., 2015), such as gait speed, stride length, fall rates, and reaction time. Otherwise, the DT tools used in the neuropsychological field involve the execution of two concurrent tasks engaging only cognitive aspects. An example is to judge the accuracy of alphanumeric equations and practice a simultaneous visual detection task under focused attention (Bier et al., 2018), as well as a single trial usually differs from a DT merely in the presentation of one or two similar stimuli (e.g., discriminating animals and/or celestial bodies by pressing a left or right button, respectively) (Lussier et al., 2012, 2021). Additional problems in the DT paradigm concern the choice of cognitive exercises. During traditional clinical performance the number of tasks that can be completed while performing a motor exercise, like walking, is restricted, especially given the small amount of time it takes to walk a few meters. Recall of words, serial subtraction, and auditory Stroop tests are common examples of activities employed (Al-Yahya et al., 2011; Raffegeau et al., 2019), although walking in a real context requires a great amount of resources such as visual attention, searching strategies, and processing of what is observed. Consequently, operating in a lifelike way is important, considering body and mind as intertwined aspects that could reflect the individual as a whole between mind and body.

WHERE COULD THIS INTEGRATION TAKE PLACE?

Both neuropsychological assessment and training are usually provided in a paper-and-pencil format and they are administered in an isolated and non-ecological setting (Gates and Valenzuela, 2010). They consist of specific domain exercises, such as categorization, classification, etc. addressing a specific ability (memory for proper names, object location, etc.) (Chaytor and Schmitter-Edgecombe, 2003; Wenisch et al., 2007; Gates and Valenzuela, 2010). Computer-based programs are progressively replacing traditional exercises since they allow for multi-modal and multi-domain training, which appears to be a key predictor of functional efficacy (Gates and Valenzuela, 2010; Hu et al., 2021). They also permit personalized intervention in a controlled way by modulating the difficulty level on the individual's baseline, gradually increasing the difficulty, and tracking results tailored

to the patient's needs. Moreover, computer-based interventions enable the real-time monitoring of cognitive performance as well as the standardization of interventions. The use of commercial packages to improve cognition or physical abilities is extremely prevalent, the Nintendo Brain training and the Brain Age software are some examples (Nouchi et al., 2012; Clemenson and Stark, 2015; Hu et al., 2021), as well as games from the Wii Fit software package (e.g., yoga, soccer, ski jump, tennis) (Franco et al., 2012; Rendon et al., 2012; Bieryla and Dold, 2013).

A further promising step could be offered by novel techniques based on immersive simulations of daily settings, i.e., Virtual Reality (VR). This is a technology that uses 3D computer-generated environments to replicate lifelike situations in an ecological, safe, and controlled situation. The fact that users behave in Virtual Environments (VE) in the same way they would in the real world distinguishes VR from other types of media (Slater and Sanchez-Vives, 2016). This illusion depends on some crucial features such as the feeling of immersion (i.e., the number of senses stimulated, valid actions that are possible within the system, and the reality's similarity of the stimuli) (Slater, 2009; Cipresso et al., 2018), the sense of presence within the environment (i.e., the feeling of "being there" inside the virtual scenes and tunes its activity accordingly within it) (Riva and Mantovani, 2014; Cipresso et al., 2018; Riva, 2018), and the interaction with objects, which allow to experience ourself as active agents (Sundar et al., 2010; Cipresso et al., 2018; Kim et al., 2019). VR also guarantees a realistic experience through multisensorial displays (e.g., visual, auditory, kinesthetic) and tracking devices that detect any movement and deliver the recorded data to the visualization system. This information about the individual is useful for a real-time update of the VE. Three categories of VR could be identified, based on the degree of immersion and interaction with the VE: non-immersive, semi-immersive, and fully immersive systems (Cipresso et al., 2018; Tuena et al., 2020). The most immersive system enables high active interaction and immersion using a head-mounted display or the Cave Automatic Virtual Environment (CAVE). They can provide a high sense of presence also by isolating individuals, facilitating natural interactions and exchanges that are similar to those found in everyday life (Riva and Mantovani, 2014; Riva, 2018). VR also provides further advantages in terms of integration since it offers the possibility to handle different devices to perform multiple tasks and collect various data. For instance, it is not only possible to cycle or walk with a dedicated bicycle or treadmill while the participant performs a second task in the VE, but also to collect behavioral data during activities using specific sensors. Thus, a subject could perform a cognitive task while cycling for example, and simultaneously digital biomarkers could be gathered, tracing what happens as it would be in daily activity. Digital biomarkers are objective and quantifiable patient data related to the behavior experienced in the VE (i.e., kinematics) and the physiological states associated with it, such as electroencephalogram (EEG), electrocardiogram (ECG), blood volume pulse (BVP), and respiration signal (RSP), detected by biosensor and medical signals. Through these variables clinicians could derive important information about neurophysiological changes (Hausdorff, 2007; Gates and Valenzuela, 2010), alterations in brain volume in specific areas (Tian et al., 2017), and consequent cognitive decline. In fact, particular patterns of gait kinematics such as slower gait, longer stride time, and more stride-to-stride variability are closely related to cognitive impairment and motor risks, with also a strong association between gait and heart rate dynamics (Hausdorff, 2007). Investigating these aspects in conjunction with cognitive requests may help to determine some clinical crucial aspects like disease severity, medication utility, and objectively document improvements in response to therapeutic interventions, above and beyond what can be gleaned from traditional measures.

HOW COULD THIS INTEGRATION TAKE PLACE?

The massive amount of information extrapolated from a patient during her/his diagnosis and treatment process needs a similarly powerful technology to process it and convert it into an output intelligible for both clinicians and patients. Machine Learning (ML) may be the most appropriate technique for managing a large volume of different, complex, and extensive data. Artificial Intelligence (AI) is a computer science field that performs activities capable of replicating human performance, such as learning to understand complex data, which is a process that requires human intelligence (Bawack et al., 2019; Graham et al., 2019; Wang, 2019). ML algorithms can discover and predict data trends and patterns by building on existing information and highlighting unexpected relationships between variables, without requiring *a priori* hypotheses about their relationships (Graham et al., 2019). Thus, ML is associated with the study and construction of systems that can learn on their own rather than following instructions. This "learning by processing" approach generates increasingly accurate predictive models for diagnosis, individual prognosis, and risk estimation (Vieira et al., 2017; Dwyer et al., 2018; Facal et al., 2019). The use of ML in healthcare has been divided into two categories: supervised (SL) and unsupervised (UL) techniques. To decide which feature best predicts the pre-labeled data, SL uses both pre-labeled data (e.g., cognitive impairment vs. healthy participants) and extra features acquired from clinical or neuroimaging sources (Dwyer et al., 2018; Graham et al., 2020). UL techniques, instead, sets unlabeled and unstructured data, e.g., clinical notes, as a starting point to seek relationships or patterns and to learn general representations that enable the automatic extraction of information when building predictors (Miotto et al., 2016; Dwyer et al., 2018; Graham et al., 2020). The great opportunity offered by AI is to combine a great number of various data, thus cognitive and motor indices may be integrated with digital biomarkers detected in the VE, offering more reliable and predictive neurophysiological results compared to the classic paper and pencil tests. AI could optimize individual treatment strategies by applying ML techniques, helping the transition to personalized, effective, and engaging medicine, built on the individual patient's needs. In the literature, ML is linked to diagnostic screening tools with subsequent analysis of the progression of the disease

(Tack, 2019; Cavedoni et al., 2020; Charvatova et al., 2020). An innovative approach is applying ML to rehabilitation, allowing precise predictions on which motor or cognitive parameters are more predictive for the future maintenance of any improvements obtained. Treatment Effect Prediction (TEP) is significant in disease management because it ensures that patients receive the expected clinical outcomes after undergoing specialized and complex treatments based on their unique clinical condition. However, there are still very few studies that analyze the potential of TEP (Shi et al., 2019; Chu et al., 2020; Wilfling et al., 2020); the latter are in the pharmacological field and use ML to predict possible side effects. Therefore, ML can augment the ability of healthcare providers to improve patient care, deliver accurate diagnoses, optimize treatment plans, inform decisions or allocate resources within health systems: precision medicine is a revolutionary approach already present in the field of pharmacology (Wilfling et al., 2020).

A PROPOSED CLINICAL APPLICATION

We suggest that VR would be the best scenario for integrating the two crucial aspects guiding daily life and those important behavioral characteristics and physiological activation. These features may be examined systematically and accurately using ML in conjunction with neuropsychological and clinical data. This comprehensive strategy could allow for the rehabilitation and the monitoring of all factors associated with aging decline. In this section, we will propose an example of this new integrated approach in the rehabilitation field.

The training will be carried out on elderly people with frailty syndrome and cognitive impairment. Participants will be randomly allocated to one of the two conditions: experimental and Treatment as Usual (TAU). Both trainings will be delivered throughout ten one-hour sessions. In TAU condition participants will complete both a paper and pencil neuropsychological rehabilitation and classic motor exercises. The experimental condition consists instead of an active VR dual-task played in the CAVE wearing 3-D glasses at IRCCS Istituto Auxologico Italiano. This immersive technology includes cameras and several head-tracked, as well as a variety of physiological and motion metrics for quantifying embodiment and movements in the VE. The protocol will be drawn based on existing tools (Pedroli et al., 2018, 2019a,b), involving different dual-task exercises: the Positive Bike, Rocks, and the Supermarket.

a. In the Positive Bike exercise, patients will use a stationary bike placed inside the CAVE. They had to maintain a constant cycling speed (motor task) and distinguish target objects from distractors (cognitive task). The exercises parameters (e.g., time between targets presentation, the target to select, bike velocity) will be decided by the therapist in each session.
b. The Rocks is designed to improve balance by immersing patients in a VE that resembles a straight road and requiring them to avoid rocks (motor task). We will add a cognitive part to this exercise: subjects will have to announce their direction while moving (i.e., if rocks move to the right, they'll say "right," and if they move to the left, they'll say "left").
c. The Supermarket exercise involves executive functions training in a virtual supermarket where patients will use an X-Box controller to move around the store and purchase many things while adhering to strict regulations. There will be ten distinct activities to choose from, each with a different level of difficulty. While shopping, we will add a motor task consisting of a walk-in place with a metronome.

During both training conditions, digital biomarkers will be collected. At the beginning of each session, all participants will wear kinematic motion detectors to detect specific bodily movements, such as velocity and stride variability. They will be attached with an elastic belt and a velcro closure to a participant's trunk and bilateral thigh and shank. Additionally, they will wear a particular h-Health smartwatch to measure heart rate variability and a chest strip sensor that will collect cardiovascular and respiratory activity. **Figure 1** shows a schematic illustration of the technological equipment.

A neuropsychological and motor assessment will be carried out before (T0), immediately after (T1), three (T2), and six (T3) months after the end of the 10 sessions, to evaluate the long-term efficacy of the treatment. To analyze data, we will assume the use of SL techniques: the goal would be to find a link between pre-labeled data and each rehabilitation session's clinical motor and cognitive findings. (i.e., clinical history of patients, standard neuropsychological battery results, and digital biomarkers). Two different types of ML will be tried out: an ML model built with the personal data set and another ML model built with the total data set. A schematic illustration of this model is depicted in **Figure 2**. In this case, however, the intent will be to show if there is any predictive long-term maintenance improvement. This would allow researchers to determine whether VR dual-task focuses on maximizing the positive effects of rehabilitation. Indeed, ML is the perfect tool, which allows analyzing the large amount of several data available: a mix of the history of the patient, results of the traditional neuropsychological test, and digital biomarkers. ML will also support the multidimensional and integrated approach aimed to use rehabilitation results and data of patients' clinical history as the input data to train a strategic dataset that in future could discriminate new cases. Particularly, using different algorithms, (e.g., Logistic Regression, Naive Bayes, Random Forest) ML would be able to track the progress of the rehabilitation of patients, its effects, and effectiveness based on the clinical history, neuropsychological tests, and parameters of digital biomarkers measured.

DISCUSSION

Aging may be a particular source of vulnerability in frailty individuals and older adults with some cognitive difficulty which may blur the border between normal and pathological aging, increasing the likelihood of chronic disease (Lutz et al., 2008). Thus, early interventions are crucial for preserving cognitive

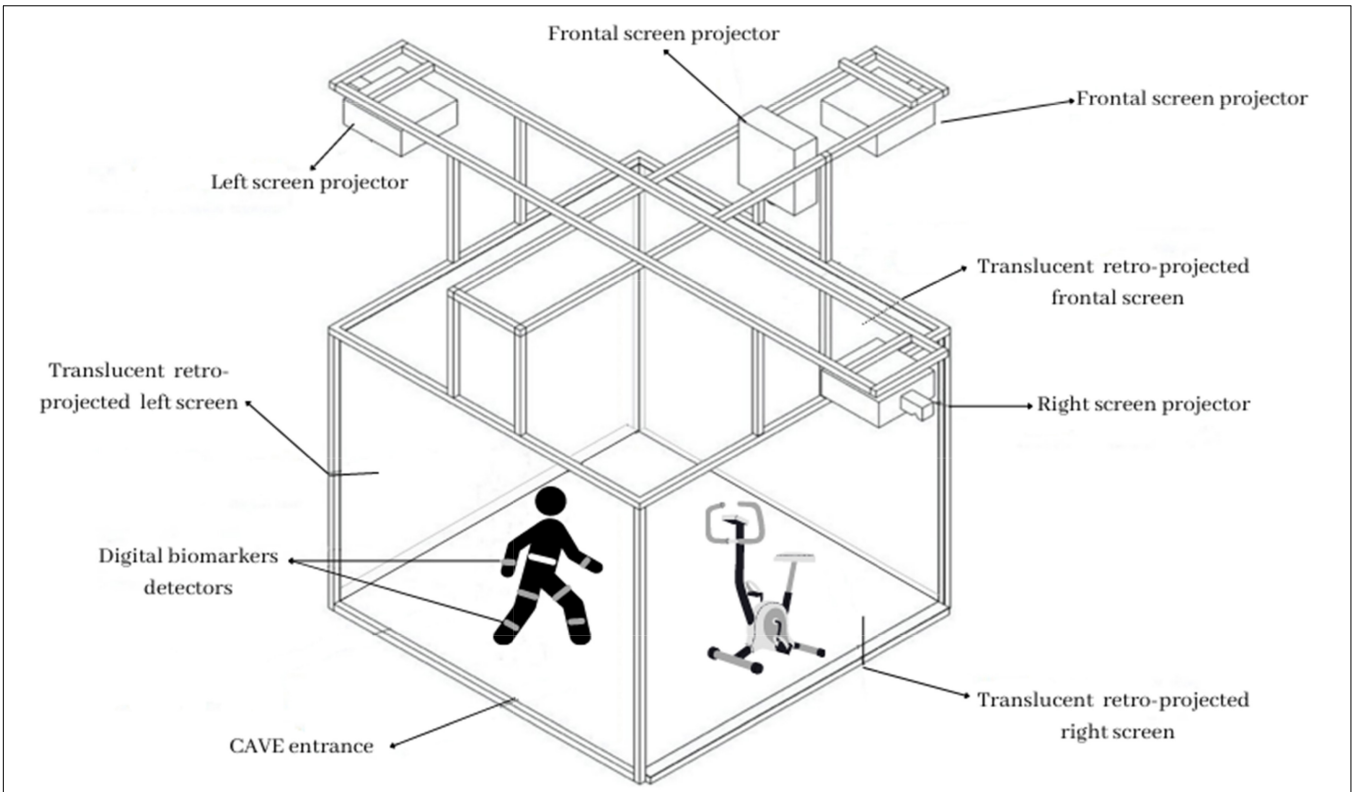


FIGURE 1 | Subject's digital biomarkers detected within a VR environment (CAVE) while performing DT.

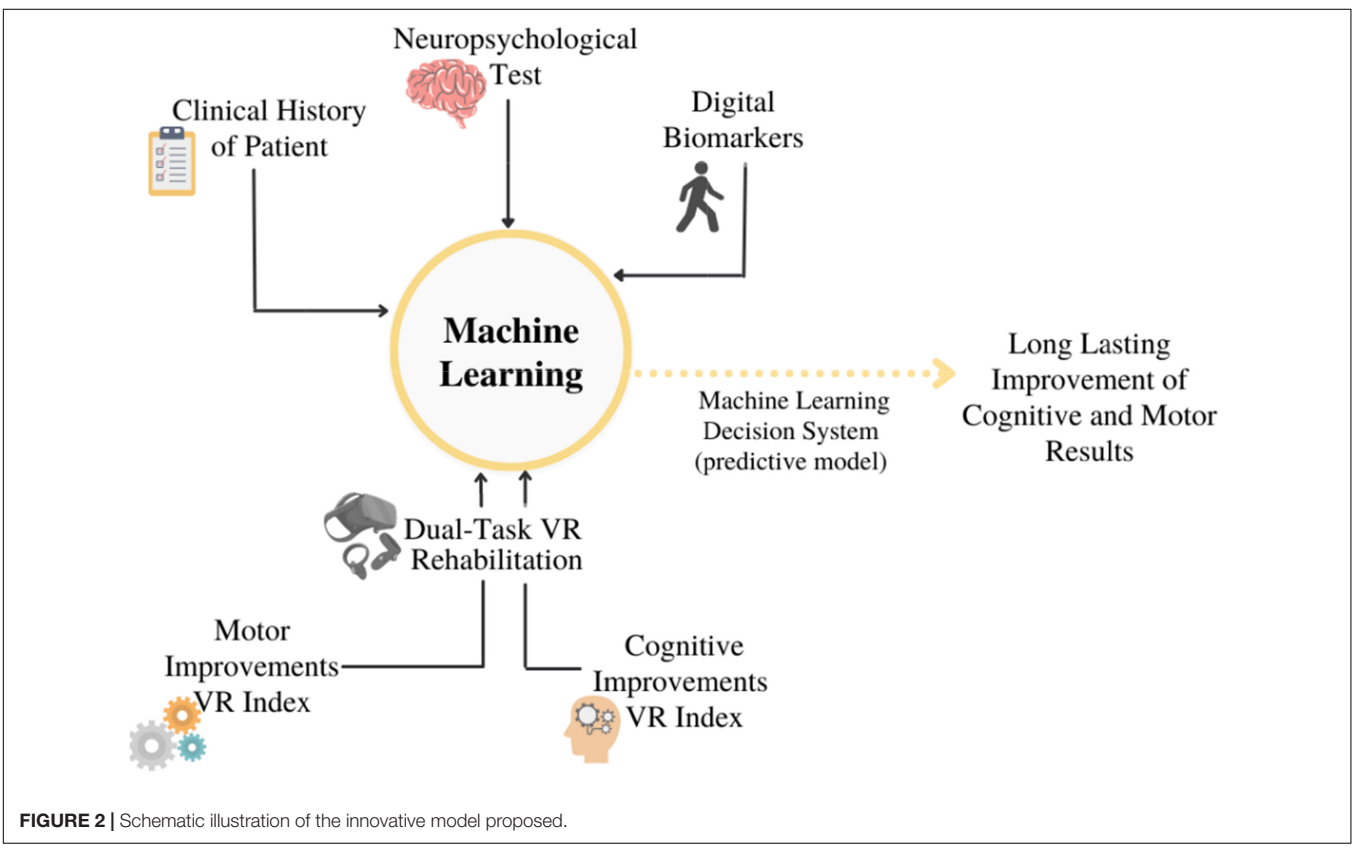


FIGURE 2 | Schematic illustration of the innovative model proposed.

functioning and, as far as possible, slowing the progression of dementia and physical impairment. When the human system is functioning properly, a multi-level integrated control system combines information from the brain with feedback from visual and proprioceptive sensors to produce, control, and modify motor patterns in response to environmental changes (Hausdorff, 2007). In this panorama, the combination of dual-task techniques with a fully immersive technology would presumably help in providing more precise results in terms of assessment. This allows for identifying real lifelike problems, and consequently, treating them with better outcomes compared to a vertical intervention because it implies the joint presence of neuropsychological and motor abilities within a realistic experience. Given the potential of VR, as well as fast feedback and repetitive practice, sophisticated cognitive processes would be facilitated by enjoyment and attractiveness, which might help motivate and engage users (Tuena et al., 2020). Patients would have to employ their attentional, mnemonic, planning, adaptability, and navigational skills to perform the virtual assignments. Thus, it is reasonable to assume that a multisession and multimodal intervention will help with the transfer of these skills to real-life daily tasks. These expected results are consistent with prior research, which has shown that performing two tasks at the same time has greater positive benefits than performing one activity at a time or sequential training (Tait et al., 2017). Further, DT is expected to increase executive functions, locomotion, and daily complicated activities in general (Mirelman et al., 2011; Wang et al., 2015; Liao et al., 2019). On these bases, the current perspective seeks to propose an innovative approach that incorporates different components while utilizing new technical devices considering both motor and cognitive baseline and improvements. Indeed, VR potential lies in duplicating lifelike behaviors (such as cycling in a park and shopping) in an ecological and controlled setting. Defining problems in aging that are as similar as in daily life allows clinicians to provide individual training that could track the specific needs of the patient that could be treated in an ecological way instead of standard interventions generally proposed in clinical practice. However, one of the problems of rehabilitation is maintaining improvements; hence, it is necessary to assess whether and how improvements will be maintained, taking into account the diversity of the results (Pashler et al., 2001; Verghese et al., 2007; Schaefer et al., 2015). To address this issue, the present perspective proposes a predictive model for detecting specific patterns of cognitive and motor impairment in aging populations and predicting outcomes. Further, digital biomarkers could also facilitate an advanced outcomes prediction and personalization of the therapeutic approach, guaranteeing customization and effective indicators of biologic processes or responses to therapeutic intervention. Indeed, sensor-based gait variability parameters were identified as clinically most relevant digital biomarkers for gait impairment (Hausdorff, 2007; Cavedoni et al., 2020; Gaßner et al., 2020) and a strong association between gait and heart rate dynamics have been identified (Hausdorff, 2007). Investigating these aspects in conjunction with cognitive requests may help to determine some clinical crucial aspects such as disease severity, medication utility, and to objectively document improvements in response to therapeutic interventions, as well as specific parameters like the difficulty and sessions, useful for effective personalized training. AI learning by processing may enable for selection of the most predictive variables, which will subsequently be empirically verified *via* follow-up; it will be also improved, based on the historical data of patients, results of neuropsychological traditional tests, and digital biomarkers measured. Measures from the initial evaluation can be compared to ML clinical data and cognitive-motor characteristics to see if and how their baseline condition influenced and predicted the outcome of rehabilitation. These integrations of different data methodologies and techniques may improve the treatment's reliability in terms of better prognostic indexes and individualized training. Clinicians might create efficient and individualized training: they could choose the best therapy in terms of intensity or number of sessions, for example, they could collect information easily about the possible future directions of the patients in terms of treatment and cognitive/motor development and saving patients' time to choose appropriate intervention in the assessment phase, by providing training quickly and efficiently. This approach crosses standardized therapy provided usually in a pre-structured way. Specific and well-timed interventions are essential to retain cognitive functioning and slow the progression of dementia and physical deterioration. For instance, early detection of gait deficits allows the early revelation not only of motor risk but also cognitive deficit, with the possibility to implement timely intervention (Lunardini et al., 2021). Further, in some cases, it could be more effective to use a vertical therapy than a traditional dual-task or a VR dual-task, allowing for long-term improvement in rehabilitative outcomes.

The integrative model of rehabilitation proposed, however, could be useful also in the assessment since data may provide the starting point for future intervention and to evaluate the person's progression, considering assessment and training as a continuum in the patient healthcare. This approach demonstrates how research and clinical practice are intertwined: research supplies the instruments, while practitioners provide critical data to confirm research creating a closed-loop system that allows researchers and clinicians to effectively interact with each other.

In summary, while much of the literature has focused on different methodologies addressed to the needs of the aging population, little was done considering the individual not as a circumscribed problem, but as a whole single person composed of mind and body at the same time. Furthermore, no one has looked at it and how some predicted characteristics are linked to the patient's fragility at the outset, opening up new avenues. In this paper, cognition meets gait in a dual-task approach that considers the individual in all her/his whole aspects. On one hand, VR provides the physical space in which all data can be collected with a high degree of ecological validity. On the other hand, ML provides the best procedures aimed to improve quality of life by lowering healthcare costs and hospitalization rates, hence expanding primary, and secondary prevention options, in the perspective of personalized medicine.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

FBr, EP, and PC developed the new model integrating virtual reality and machine learning. FBo supervised the sections of machine learning. VM supervised the section regarding virtual reality. FBr wrote the manuscript under the final supervision of EP and PC. GR and MS-B provided the required revisions. All authors have approved the final version of the manuscript.

FUNDING

This research was funded by the Italian Ministry of Health (POSTECH: 39C801_2018).

REFERENCES

Al-Yahya, E., Dawes, H., Smith, L., Dennis, A., Howells, K., and Cockburn, J. (2011). Cognitive motor interference while walking: a systematic review and meta-analysis. *Neurosci. Biobehav .Rev.* 35, 715–728. doi: 10.1016/j.neubiorev.2010.08.008

Bawack, R, Fosso Wamba, S., and Carillo, K. (2019). *Artificial Intelligence in Practice: Implications for Information Systems Research*. Mexico: Americas Conference on Information Systems

Bier, B., Ouellet, É., and Belleville, S. (2018). Computerized attentional training and transfer with virtual reality: effect of age and training type. *Neuropsychology* 32, 597–614. doi: 10.1037/neu0000417

Bieryla, K. A., and Dold, N. M. (2013). Feasibility of Wii Fit training to improve clinical measures of balance in older adults. *Clin. Interv. Aging* 8, 775–781.

Cavedoni, S., Chirico, A., Pedroli, E., Cipresso, P., and Riva, G. (2020). Digital Biomarkers for the Early Detection of Mild Cognitive Impairment: artificial Intelligence Meets Virtual Reality. *Front. Hum. Neurosci.* 14:245. doi: 10.3389/fnhum.2020.00245

Charvatova, H., Procházka, A., and Vyšata, O. (2020). Motion Assessment for Accelerometric and Heart Rate Cycling Data Analysis. *Sensors* 20:1523. doi: 10.3390/s20051523

Chaytor, N., and Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. *Neuropsychology* 13, 181–197. doi: 10.1023/b:nerv.0000009483.91468.fb

Chu, J., Dong, W., Wang, J., He, K., and Huang, Z. (2020). Treatment effect prediction with adversarial deep learning using electronic health records. *BMC Med. Inform. Decis. Mak.* 20:139. doi: 10.1186/s12911-020-01151-9

Cipresso, P., Giglioli, I. A. C., Raya, M. A., and Riva, G. (2018). The Past, Present, and Future of Virtual and Augmented Reality Research: a Network and Cluster Analysis of the Literature. *Front. Psychol.* 9:2086. doi: 10.3389/fpsyg.2018.02086

Clemenson, G. D., and Stark, C. E. L. (2015). Virtual environmental enrichment through video games improves hippocampal-associated memory. *J. Neurosci.* 35, 16116–16125. doi: 10.1523/JNEUROSCI.2580-15.2015

de Morais Fabrício, D., Chagas, M. H. N., and Diniz, B. S. (2020). Frailty and cognitive decline. *Transl. Res.* 221, 58–64.

de Vries, N. M., Staal, J. B., van Ravensberg, C. D., Hobbelen, J. S., and Olde Rikkert, M. G. (2011). Nijhuis-van der Sanden MW. Outcome instruments to measure frailty: a systematic review. *Ageing Res. Rev.* 10, 104–114. doi: 10.1016/j.arr.2010.09.001

Deutsch, J. E., Myslinski, M. J., Kafri, M., Ranky, R., Sivak, M., Mavroidis, C., et al. (2013). Feasibility of virtual reality augmented cycling for health promotion of people poststroke. *J. Neurol. Phys. Ther.* 37, 118–124. doi: 10.1097/NPT.0b013e3182a0a078

Dorfman, M., Herman, T., Brozgol, M., Shema, S., Weiss, A., Hausdorff, J. M., et al. (2014). Dual-task training on a treadmill to improve gait and cognitive function in elderly idiopathic fallers. *J. Neurol. Phys. Ther.* 38, 246–253. doi: 10.1097/NPT.0000000000000057

Dwyer, D. B., Falkai, P., and Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annu. Rev. Clin. Psychol.* 14, 91–118.

Facal, D., Valladares-Rodriguez, S., Lojo-Seoane, C., Pereiro, A. X., Anido-Rifon, L., and Juncos-Rabadán, O. (2019). Machine learning approaches to studying the role of cognitive reserve in conversion from mild cognitive impairment to dementia. *Int. J. Geriatr. Psychiatry* 34, 941–949. doi: 10.1002/gps.5090

Franco, J. R., Jacobs, K., Inzerillo, C., and Kluzik, J. (2012). The effect of the Nintendo Wii Fit and exercise in improving balance and quality of life in community dwelling elders. *Technol. Health Care* 20, 95–115. doi: 10.3233/THC-2011-0661

Freitag, F., Brucki, S. M. D., Barbosa, A. F., Chen, J., Souza, C. O., Valente, D. F., et al. (2019). Is virtual reality beneficial for dual-task gait training in patients with Parkinson's disease? A systematic review. *Dement. Neuropsychol.* 13, 259–267. doi: 10.1590/1980-57642018dn13-030002

Fried, L. P., Ferrucci, L., Darer, J., Williamson, J. D., and Anderson, G. (2004). Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *J. Gerontol. A Biol. Sci. Med. Sci.* 59, 255–263. doi: 10.1093/gerona/59.3.m255

Gates, N., and Valenzuela, M. (2010). Cognitive exercise and its role in cognitive function in older adults. *Curr. Psychiatry Rep.* 12, 20–27. doi: 10.1007/s11920-009-0085-y

Gaßner, H., Jensen, D., Marxreiter, F., Kletsch, A., Bohlen, S., Schubert, R., et al. (2020). Gait variability as digital biomarker of disease severity in Huntington's disease. *J. Neurol.* 267, 1594–1601. doi: 10.1007/s00415-020-09725-3

Gobbens, R. J., Luijkx, K. G., Wijnen-Sponselee, M. T., and Schols, J. M. (2010). Toward a conceptual definition of frail community dwelling older people. *Nurs. Outlook* 58, 76–86. doi: 10.1016/j.outlook.2009.09.005

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., et al. (2019). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Curr. Psychiatry Rep.* 21:116. doi: 10.1007/s11920-019-1094-0

Graham, S. A., Lee, E. E., Jeste, D. V., Van Patten, R., Twamley, E. W., Nebeker, C., et al. (2020). Artificial intelligence approaches to predicting and detecting cognitive decline in older adults: a conceptual review. *Psychiatry Res.* 284:112732. doi: 10.1016/j.psychres.2019.112732

Hausdorff, J. M. (2007). (2007). Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking. *Eur. Workshop Mov. Sci.* 26, 555–589. doi: 10.1016/j.humov.2007.05.003

Hu, M., Wu, X., Shu, X., Hu, H., Chen, Q., Peng, L., et al. (2021). Effects of computerised cognitive training on cognitive impairment: a meta-analysis. *J. Neurol.* 268, 1680–1688. doi: 10.1007/s00415-019-09522-7

Kim, O., Pang, Y., and Kim, J. H. (2019). The effectiveness of virtual reality for people with mild cognitive impairment or dementia: a meta-analysis. *BMC Psychiatry* 19:219. doi: 10.1186/s12888-019-2180-x

Kojima, G., Iliffe, S., and Walters, K. (2018). Frailty Index as a Predictor of Mortality: A Systematic Review and meta-analysis. *Age Ageing.* 47, 193–200.

Lauenroth, A., Ioannidis, A. E., and Teichmann, B. (2016). Influence of combined physical and cognitive training on cognition: a systematic review. *BMC Geriatr.* 16:141. doi: 10.1186/s12877-016-0315-1

Liao, Y. Y., Chen, I. H., Lin, Y. J., Chen, Y., and Hsu, W. C. (2019). Effects of Virtual Reality-Based Physical and Cognitive Training on Executive Function and Dual-Task Gait Performance in Older Adults With Mild Cognitive Impairment: a Randomized Control Trial. *Front. Aging Neurosci.* 11:162. doi: 10.3389/fnagi.2019.00162

Lunardini, F., Malavolti, M., Pedrocchi, A. L. G., Borghese, N. A., and Ferrante, S. (2021). A mobile app to transparently distinguish single- from dual-task walking for the ecological monitoring of age-related changes in daily-life gait. *Gait Posture* 86, 27–32. doi: 10.1016/j.gaitpost.2021.02.028

Lussier, M., Gagnon, C., and Bherer, L. (2012). An investigation of response and stimulus modality transfer effects after dual-task training in younger and older. *Front. Hum. Neurosci.* 6:129. doi: 10.3389/fnhum.2012.00129

Lussier, M., Saillant, K., Vrinceanu, T., Hudon, C., and Bherer, L. (2021). Normative Data for a Tablet-Based Dual-Task Assessment in Healthy Older Adults. *Arch. Clin. Neuropsychol.* 36, 1316–1325. doi: 10.1093/arclin/acaa121

Lutz, W., Sanderson, W., and Scherbov, S. (2008). The coming acceleration of global population ageing. *Nature* 451, 716–719.

Miotto, R., Li, L., Kidd, B. A., and Dudley, J. T. (2016). Deep Patient: an Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Sci. Rep.* 6:26094. doi: 10.1038/srep26094

Mirelman, A., Maidan, I., Herman, T., Deutsch, J. E., Giladi, N., and Hausdorff, J. M. (2011). Virtual reality for gait training: can it induce motor learning to enhance complex walking and reduce fall risk in patients with Parkinson's disease? *J. Gerontol. A Biol. Sci. Med. Sci.* 66, 234–240. doi: 10.1093/gerona/glq201

Montero-Odasso, M., and Speechley, M. (2018). Falls in Cognitively Impaired Older Adults: implications for Risk Assessment And Prevention. *J. Am. Geriatr. Soc.* 66, 367–375.

Nouchi, R., Taki, Y., Takeuchi, H., Hashizume, H., Akitsuki, Y., Shigemune, Y., et al. (2012). Brain training game improves executive functions and processing speed in the elderly: a randomized controlled trial. *PLoS One* 7:e29676. doi: 10.1371/journal.pone.0029676

Panza, F., Seripa, D., Solfrizzi, V., Tortelli, R., Greco, A., Pilotto, A., et al. (2015). Targeting Cognitive Frailty: clinical and Neurobiological Roadmap for a Single Complex Phenotype. *J. Alzheimers Dis.* 47, 793–813. doi: 10.3233/JAD-150358

Pashler, H., Johnston, J. C., and Ruthruff, E. (2001). Attention and performance. *Annu. Rev. Psychol.* 52, 629–651.

Pedroli, E., Cipresso, P., Greci, L., Arlati, S., Boilini, L., Stefanelli, L., et al. (2019a). An Immersive Motor Protocol for Frailty Rehabilitation. *Front. Neurol.* 10:1078. doi: 10.3389/fneur.2019.01078

Pedroli, E., Cipresso, P., Serino, S., Toti, M., Goulen, K., Grigioni, M., et al. (2019b). "Beyond Cognitive Rehabilitation: immersive but Noninvasive Treatment for Elderly," in *Pervasive Computing Paradigms for Mental Health. MindCare 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, eds P. Cipresso, S. Serino, and D. Villani (Cham: Springer), doi: 10.1007/978-3-030-25872-6_22

Pedroli, E., Greci, L., Colombo, D., Serino, S., Cipresso, P., Arlati, S., et al. (2018). Characteristics, Usability, and Users Experience of a System Combining Cognitive and Physical Therapy in a Virtual Environment: positive Bike. *Sensors* 18:2343. doi: 10.3390/s18072343

Raffegeau, T. E., Krehbiel, L. M., Kang, N., Thijs, F. J., Altmann, L. J. P., Cauraugh, J. H., et al. (2019). A meta-analysis: parkinson's disease and dual-task walking. *Parkinsonism Relat. Disord.* 62, 28–35. doi: 10.1016/j.parkreldis.2018.12.012

Rendon, A. A., Lohman, E. B., Thorpe, D., Johnson, E. G., Medina, E., and Bradley, B. (2012). The effect of virtual reality gaming on dynamic balance in older adults. *Age Ageing* 41, 549–552. doi: 10.1093/ageing/afs053

Riva, G. (2018). The neuroscience of body memory: from the self through the space to the others. *Cortex* 104, 241–260. doi: 10.1016/j.cortex.2017.07.013

Riva, G., and Mantovani, F. (2014). "Extending the Self through the Tools and the Others: a General Framework for Presence and Social Presence in Mediated Interactions," in *Interacting with Presence: HCI and the Sense of Presence in Computer-Mediated Environments*, eds G. Riva, J. Waterworth, and D. Murray (Berlin: De Gruyter), 9–31.

Rockwood, K. (2005). What would make a definition of frailty successful? *Age Ageing* 34, 432–434.

Schaefer, S., Schellenbach, M., Lindenberger, U., and Woollacott, M. (2015). Walking in high-risk settings: do older adults still prioritize gait when distracted by a cognitive task? *Exp. Brain Res.* 233, 79–88. doi: 10.1007/s00221-014-4093-8

Shi, C., Blei, D., and Veitch, V. (2019). Adapting neural networks for the estimation of treatment effects. *Adv. Neural. Inf. Process. Syst.* 32, 2503–2513.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 364, 3549–3557. doi: 10.1098/rstb.2009.0138

Slater, M., and Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality. *Front. Robot. AI* 3:74. doi: 10.3389/frobt.2016.00074

Speechley, M., and Tinetti, M. (1991). Falls and injuries in frail and vigorous community elderly persons. *J. Am. Geriatr. Soc.* 39, 46–52.

Sundar, S. S., Xu, Q., and Bellur, S. (2010). "Designing Interactivity in Media Interfaces: a Communications Perspective," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (New York, NY: Association for 470 Computing Machinery), 2247–2256.

Tack, C. (2019). Artificial intelligence and machine learning| applications in musculoskeletal physiotherapy. *Musculoskelet. Sci. Pract.* 39:164169.

Tait, J. L., Duckham, R. L., Milte, C. M., Main, L. C., and Daly, R. M. (2017). Influence of Sequential vs. Simultaneous Dual-Task Exercise Training on Cognitive Function in Older Adults. *Front. Aging Neurosci.* 9:368. doi: 10.3389/fnagi.2017.00368

Tian, Q., Chastan, N., Bair, W. N., Resnick, S. M., Ferrucci, L., and Studenski, S. A. (2017). The brain map of gait variability in aging, cognitive impairment and dementia-A systematic review. *Neurosci. Biobehav. Rev.* 74, 149–162. doi: 10.1016/j.neubiorev.2017.01.020

Tuena, C., Pedroli, E., Trimarchi, P. D., Gallucci, A., Chiappini, M., Goulene, K., et al. (2020). Usability Issues of Clinical and Research Applications of Virtual Reality in Older People: a Systematic Review. *Front. Hum. Neurosci.* 14:93. doi: 10.3389/fnhum.2020.00093

Verghese, J., Wang, C., Lipton, R. B., Holtzer, R., and Xue, X. (2007). Quantitative gait dysfunction and risk of cognitive decline and dementia. *J. Neurol. Neurosurg. Psychiatry* 78, 929–935. doi: 10.1136/jnnp.2006.106914

Vieira, S., Pinaya, W. H. L., and Mechelli, A. (2017). Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. *Neurosci. Biobehav. Rev.* 74, 58–75. doi: 10.1016/j.neubiorev.2017.01.002

Wajda, D. A., Mirelman, A., Hausdorff, J. M., and Sosnoff, J. J. (2017). Intervention modalities for targeting cognitive-motor interference in individuals with neurodegenerative disease: a systematic review. *Expert. Rev. Neurother.* 17, 251–261. doi: 10.1080/14737175.2016.1227704

Wang, P. (2019). On Defining Artificial Intelligence. *J. Artif. Gen. Intell* 10, 1–37.

Wang, X., Pi, Y., Chen, P., Liu, Y., Wang, R., and Chan, C. (2015). Cognitive motor interference for preventing falls in older adults: a systematic review and meta-analysis of randomised controlled trials. *Age Ageing* 44, 205–212. doi: 10.1093/ageing/afu175

Wenisch, E., Cantegreil-Kallen, I., De Rotrou, J., Garrigue, P., Moulin, F., Batouche, F., et al. (2007). Cognitive stimulation intervention for elders with mild cognitive impairment compared with normal aged subjects: preliminary results. *Aging Clin. Exp. Res.* 19, 316–322. doi: 10.1007/BF03324708

Wilfling, D., Hinz, A., and Steinhäuser, J. (2020). Big data analysis techniques to address polypharmacy in patients – a scoping review. *BMC Fam. Pract.* 21:180. doi: 10.1186/s12875-020-01247-1

World Health Organization (2020). *Decade of healthy ageing: baseline report*. Geneva: World Health Organization

Yogev, G., Hausdroff, J. M, and Gilardi, N. (2008). The Role of Executive Function and Attention in Gait. *Mov. Disord.* 23, 329–342.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 28 July 2022
DOI 10.3389/fnagi.2022.875748

OPEN ACCESS

EDITED BY
Yang Jiang,
University of Kentucky, United States

REVIEWED BY
Amanda Selwood,
University of New South Wales, Australia
Martin Dyrba,
German Center for Neurodegeneratives (HZ), Germany

*CORRESPONDENCE
Francesca Bruni
francesca.bruni016@gmail.com

SPECIALTY SECTION
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 14 February 2022
ACCEPTED 05 July 2022
PUBLISHED 28 July 2022

CITATION
Bruni F, Mancuso V, Stramba-Badiale C, Greci L, Cavallo M, Borghesi F, Riva G, Cipresso P, Stramba-Badiale M and Pedroli E (2022) ObReco-2: Two-step validation of a tool to assess memory deficits using 360° videos. *Front. Aging Neurosci.* 14:875748. doi: 10.3389/fnagi.2022.875748

ObReco-2: Two-step validation of a tool to assess memory deficits using 360° videos

Francesca Bruni[1]*, Valentina Mancuso[1], Chiara Stramba-Badiale[2], Luca Greci[3], Marco Cavallo[1], Francesca Borghesi[2], Giuseppe Riva[2,4], Pietro Cipresso[2,5], Marco Stramba-Badiale[6] and Elisa Pedroli[1,2]

[1]Faculty of Psychology, eCampus University, Novedrate, Italy, [2]Applied Technology for Neuropsychology Lab, IRCCS Istituto Auxologico Italiano, Milan, Italy, [3]Institute of Intelligent Industrial Technologies and System for Advanced Manufacturing, Milan, Italy, [4]Human Technology Lab, Catholic University of the Sacred Heart, Milan, Italy, [5]Department of Psychology, University of Turin, Turin, Italy, [6]Department of Geriatrics and Cardiovascular Medicine, IRCCS Istituto Auxologico Italiano, Milan, Italy

Traditional neuropsychological evaluations are usually carried out using psychometric paper and pencil tests. Nevertheless, there is a continuous discussion concerning their efficacy to capture life-like abilities. The introduction of new technologies, such as Virtual Reality (VR) and 360° spherical photos and videos, has improved the ecological validity of the neuropsychological assessment. The possibility of simulating realistic environments and situations allows clinicians to evaluate patients in realistic activities. Moreover, 360° photos and videos seem to provide higher levels of graphical realism and technical user-friendliness compared to standard VR, regardless of their limitations in terms of interactivity. We developed a novel 360° tool, ObReco-2 (Object Recognition version 2), for the assessment of visual memory which simulates a daily situation in a virtual house. More precisely, patients are asked to memorize some objects that need to be moved for a relocation. After this phase, they are asked to recall them after 15 min and later to recognize them in the same environment. Here we present a first study about the usability of ObReco-2, and a second one exploring its clinical efficacy and updated usability data. We focused on Free Recall and Recognition scores, comparing the performances obtained by the participants in the standard and the 360° test. The preliminary results support the use of 360° technology for enhancing the ecological value of standard memory assessment tests.

KEYWORDS

memory, neuropsychological assessment, 360° video, virtual reality, object recognition, neuroscience

Introduction

Recently, the debate regarding the ecological validity of the measures typically employed for the assessment of cognitive domains seems to be an open-ended question in the neuropsychological field. Ecological validity refers to the degree of association

between what is observed during neuropsychological testing and real-life activities, i.e., the ability of paper and pencil tests to predict real-life functioning (Sbordone, 1996). Despite the widespread use of paper and pencil tests, their ability to predict patients' skills in real-life circumstances could be limited (Neguț et al., 2016; Rizzo and Koenig, 2017). One of the main issues is that patients may not show deficits in the clinical setting but at the same time report some difficulties in everyday situations or vice versa (Chaytor and Schmitter-Edgecombe, 2003; Mondini et al., 2016). Indeed, during the clinical evaluation patients are required to carry out various behavioral and cognitive activities in a controlled setting which may not always predict their functioning in a real-life situation. Therefore, it is worth considering the debate over the efficacy of many traditional tests assuming a more function-based approach rather than a construct-based one (Parsons, 2015; Parsons et al., 2017; Serino and Repetto, 2018). A construct-based approach starts from a solid theoretical paradigm assessing abstract constructs without an explicit interest in predicting real-life functional abilities. On the other hand, a function-based approach arises from direct observations of patients' performance in real-life contexts to guarantee a more ecological assessment (Sbordone, 1996; Parsons, 2015; Parsons et al., 2017). The Rivermead Behavioural Memory Test (RBMT) is the most well-known example of this approach for memory assessment (Wilson et al., 1989). It includes a series of daily-life tasks such as locating personal objects, remembering an appointment, recalling an itinerary, etc.

In recent years, technologies might be considered promising realities in accomplishing and improving ecological validity, sensitivity, and specificity of traditional assessment methods. Among these, virtual reality (VR) emerges as a suitable possibility in neuropsychological assessment. This technology can be employed to develop highly ecological and controlled environments resembling the real-life contexts in which patients' daily activities usually take place (Riva and Mantovani, 2014; Neguț et al., 2016; Riva et al., 2019). It thus can allow researchers and clinicians to measure cognitive and motor abilities in naturalistic environments, obtaining better prognostic indexes of real-life functioning in a safe and controlled situation. This approach has been widely used in the medical and neuropsychological field to assess and treat different pathologies such as traumatic brain injury (Aida et al., 2018; Alashram et al., 2019) and post-stroke (Saposnik and Levin, 2011; Laver et al., 2017). Moreover, it has been revealed promising for balance deficits (Allain et al., 2014) and memory impairments (Matheis et al., 2007; Ouellet et al., 2018; Serino and Repetto, 2018). More specifically, memory interventions included several tasks in which patients were required to perform some activities while navigating in the 3D environments (i.e., office and supermarket) (Matheis et al., 2007; Ouellet et al., 2018; Serino and Repetto, 2018). The employment of 360° immersive photos and videos is a growing declination of VR technology that may offer promising outcomes (Serino and Repetto, 2018; Realdon et al., 2019; Ventura et al., 2019). They are spherical videos or photos captured by an omnidirectional camera. As previously mentioned, this method has greater benefits than graphic-based VR as it can capture the real environment, providing a high level of visual realism that can increase participant engagement. Moreover, this technology is inexpensive and easy-to-use (Bohil et al., 2011). Furthermore, the user-friendly design makes 360° technologies more suitable for the assessment of patients with mild to severe impairments (Sbordone, 1996; Realdon et al., 2019) who may have some difficulties interacting with more sophisticated devices.

The present study aims to test a 360° technology for memory assessment compared to a traditional paper and pencil test included in the RBMT-III (Wilson et al., 2008; Beschin and Urbano, 2013). Based on promising results from an earlier pilot study showing the feasibility of a 360° memory assessment (Pieri et al., 2021), we improved technology using higher-level equipment to design ObReco-2 (Object Recognition version 2). Firstly, we present the results of a usability study (Study 1), and then the results of the clinical efficacy along with updated usability data (Study 2).

Study 1 (usability study)

Materials and methods

Participants

For the usability assessment, participants were enrolled among the patients and outpatients of the Department of Medical Rehabilitation of Istituto Auxologico Italiano in Milan. They were volunteers aged over 60 (without maximum age limitation), with a normal or corrected-to-normal vision. Exclusion criteria were: (i) invalidating internist, psychiatric, neurological conditions which could affect the usability of the task; (ii) cognitive impairments certifiable by a score at the Mini-Mental State Examination (MMSE) Italian version (Measso et al., 1993; Magni et al., 1996) lower than 24 points. The resulting sample included 10 participants (6 females and 4 males), with a mean age of 75.5 (SD = 5.36) and a mean of 12.3 (SD = 3.89) years of education. All the subjects' demographic data and MMSE scores are reported in **Table 1**. Before the usability session, all participants signed the informant consent. The study received ethical approval from the Ethical Committee of the Istituto Auxologico Italiano.

Materials

Files were recorded in a real environment using the Insta 360 ONE X, an omnidirectional video camera that can record spherical photos and videos with a resolution respectively of 608 × 3040 and 5.760 × 2.880 pixels. We combined all

TABLE 1 Demographic data and mini-mental state examination scores.

Descriptives			
	Years	**Education**	**MMSE**
Mean	75.5	12.3	25.8
Standard deviation	5.36	3.89	1.47
Min	68	4	23.3
Max	84	18	28.0

photos and videos into a single interactive experience, using the InstaVR software©. The result consists of an application deliverable via smartphone that may be experienced using a Cardboard, which allows the user to navigate within this immersive 360° scenario. In particular, the application was provided via an InstaVR link on the smartphone which was inserted into the Cardboard to show the environment.

Procedure

For this study participants were examined in two sessions at a maximum of 2 days apart. In the first session, the MMSE was administered to quantify the general cognitive state of the patients. The second phase of the study consisted of a usability study employing cardboard (Daydream view©). Usability is a key factor that needs to be evaluated when employing new technologies. It can be defined as the degree to which a user can utilize a given system to achieve specific goals effectively, efficiently, and satisfactorily. Usability test allows the clinicians to identify obstacles and facilitators, develop appropriate tasks for the target, define the usability criteria and test its clinical use (Tuena et al., 2020). During the usability session, all the participants were sitting on a turning chair, to freely explore the 360° virtual environments using the cardboard.

User experience measures

In the present study, the usability has been assessed using the System Usability Scale (SUS) (Brooke, 2020), the Senior Technology Acceptance Model (STAM) (Chen and Lou, 2020), the thinking aloud protocol (TAP) (Lewis, 1982), and the Independent Television Commission Sense of Presence Inventory (ITC-SOPI) to assess the cybersickness (Lessiter et al., 2001). The SUS (Brooke, 2020) is a "quick and easy-to-use" questionnaire which includes ten items describing the user's feeling concerning the interaction with the technology. For each of these answers, the participants need to define their degree of agreement using a 5-point Likert scale ranging from "Strongly Agree" to "Strongly Disagree". The final score ranges from 0 (lack of usability) to 100 (optimal usability). The STAM is a 13-items questionnaire that analyzes four components of the STAM: attitude through technologies, perception of control, anxiety related to technologies, and general health conditions (Chen and Lou, 2020). The TAP (Lewis, 1982) is a qualitative technique that is generally administered to test the usability of new technology. Subjects are asked to express their opinion regarding the technology employment and criticism while performing the task. The observer, on the other hand, is asked to take notes of participants' observations and concerns without attempting to interpret their actions and words. All the verbalizations are transcribed and analyzed to develop the formal usability report. The ITC-SOPI (Lessiter et al., 2001) is a questionnaire that includes 44 items addressing the individual's feelings after the VR experience. Participants are asked to determine their degree of agreement with each of these sentences using a 5-point Likert scale ranging from "Strongly Agree" to "Strongly Disagree". The ITC-SOPI includes four subscales: Sense of Physical Space (19 items), Engagement (13 items), Ecological Validity (5 items), and Negative Effects (6 items).

ObReco-2

ObReco-2 is a 360° task aimed to assess visual memory simulating a real-life situation in a daily setting. Users are immersed in a virtual living room, in which they are required to encode and then recall some target objects that have been relocated, as described in Bruni et al. (2022). The virtual interactive experience consists of a series of different phases:

(i) *Familiarization*. Patients who wear the headset find themselves immersed in a natural 360° landscape; here they have to explore the environment. The objective is to make the patient familiar with the technology and to detect possible side effects (i.e., cybersickness).

(ii) *Encoding*. On a black screen, the participants are first given a brief explanation of the context: Marco, who is living with other roommates, must move and he had to relocate all of his possessions, thus he labels them with his name. Participants experience a household setting, such as a living room, in which they can with a first-person perspective which is the one of the experimenter. This one moves about the room highlighting the 15 target items for 3 s each and attaching a tag bearing the name "Marco" to each one **Figure 1**). In the living room, there are also 15 other objects used as distractors. In this phase, participants are instructed to name all the targets.

(iii) *Interference*. Participants are asked to take off their headsets and complete non-verbal tasks15 for minutes.

(iv) *Free recall*. They are instructed to name as many objects from the encoding phase as they can.

(v) *Recognition*. Participants had to wear the headset once again for this last section. They are instructed to explore the prior living room (**Figure 2**), discover and name the target objects among all of the previous things and an unknown set of 15 distractor objects.

FIGURE 1
A screenshot showing the task presented during the encoding phase. Here the experimenter is labeling a target object.

FIGURE 2
The panoramic photo of the room in which target objects are mixed with distractors.

Data analysis

We organized all the data collected in a Windows Excel sheet and we performed descriptive analyses of the usability questionnaires investigating users' experience with technology.

Results

The descriptive user experience (UX) measures are shown in **Table 2**.

Starting from quantitative data, the mean score of the SUS is 69.3 (SD = 18.1). According to Bangor et al. (2009) this score indicates that ObReco-2 is placed in a marginal zone between *Ok* and a *Good level* of usability as shown in **Figure 3**.

The results of the STAM scale reveal that our sample has a positive attitude toward technology (M = 6.10/10; SD = 3.7), has good control/access to technological devices (M = 7,13/10; SD = 1.98), has a medium level of technology-related anxiety (M = 5.60/10; SD = 3.07), and considers themselves in good health conditions (M = 8.58/10; SD = 1.8). As shown by the ITC-SOPI sub-scale investigating negative effects, all subjects reported minimal side effects (M = 1.87; SD = 0.90) indicating that the use of ObReco-2 did not determine dizziness and cybersickness. Qualitative results of the thinking aloud protocol are shown in **Table 3**. It is structured as follows: (i) description of the task (1st column), (ii) problems encountered by patients (2nd column), (iii) some possible solution for those problems (3rd column), and (iv) number of patients that encountered problems (4th column). Overall, patients did not encounter problems using the cardboard. However, most patients reported difficulties in the encoding exercise in which they were required to explore the room following labels and naming Marco's objects. Five patients had difficulty in exploring the environment. Four patients reported unclear images; one didn't name all the objects and four people had difficulty finding the initial correct direction to follow the 360° video. Finally, one person reported nausea.

Study 2 (usability study and clinical efficacy)

Materials and methods

Participants

For this clinical efficacy and usability study, 20 patients were enrolled at the Department of Medical Rehabilitation of Istituto Auxologico Italiano in Milan. They were volunteers aged over 55, with a normal or corrected-to-normal vision. Exclusion criteria were (i) invalidating internist, psychiatric, or neurological conditions which could affect the task; (ii) cognitive impairments difficulties certifiable by a score at the MMSE Italian version lower than 24 points (Measso et al., 1993; Magni et al., 1996). Before the session, all participants signed the informant consent. The study received ethical approval from the Ethical Committee of the Istituto Auxologico Italiano. The sample was composed by 12 females and 8 males, divided in experimental group (ObReco-2–VR) (Mean age = 68.2 years, SD = 5.45, mean education = 12, SD = 4.45, 6 females) and control group (RBMT-III—paper and pencil) (Mean age = 69.7 years, SD = 7.63, mean education = 14.6, SD = 3.84, 6 females). All the subjects' demographic data and MMSE scores are reported in **Table 4**. The two groups are comparable in age $t(18) = 0.506$, $p = 0.619$, in years of education $t(18) = 1.400$, $p = 0.179$ and MMSE $t(18) = 0.506$, $p = 0.619$.

Materials

Improving VR experience, we implemented all the previously collected files into a single interactive experience, using Unity3D©. The result consists of an interactive application deliverable through a head-mounted display (HMD), which allows the user to navigate and interact within the immersive 360° scenario. The application was downloaded and installed directly on an Oculus Quest-2© HMD to be used without any restrictions. Thanks to the most advanced

TABLE 2 Descriptives of the user experience (UX) measures.

	SUS	STAM-a	STAM-c	STAM-anx	STAM-h	ITC-sp	ITC-e	ITC-ev	ITC-ne
Mean	69.3/100	6.10/10	7.13/10	5.60/10	8.58/10	2.94/5	3.33/5	3.96/5	1.87/5
Standard deviation	18.1	3.07	1.98	3.07	1.80	1.03	0.73	0.94	0.90
Min	40.0	1.00	3.25	1.00	4.60	1.18	1.77	2.00	1.00
Max	100	9.33	10.0	10.0	11.2	4.25	4.31	4.80	3.83

For each measure there are mean and the maximum available score, standard deviation and minimum (min) and maximum (max) score reported by participants.
SUS, System Usability Scale; CSQ; STAM-a, attitude through technologies subscale; STAM-c, Senior Technology Acceptance Model perception of control subscale; STAM-anx, Senior Technology Acceptance Model anxiety related to technologies subscale; STAM-h, Senior Technology Acceptance Mode health conditions subscale; ITC-sp, Independent Television Commission Sense of Presence Inventory-Sense of Physical Space subscale; ITC-e, Engagement subscale; ITC-ev, Ecological Validity subscale; ITC-ne, Negative Effects subscale.

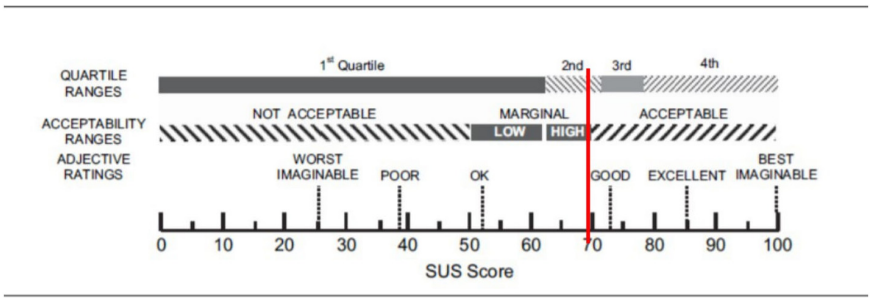


FIGURE 3
Graphical representation of the interpretation of system usability scale (SUS). The vertical line shows the position of the SUS mean score (69.3) obtained in study 1 according to the rating comparison scale provided by Bangor et al. (2009).

functionalities, participants can have major interactivity with the environment, without the continuous intervention of the experimenter. In the first usability study, subjects were limited in their interaction with the environment; different links were provided to them, corresponding to the different parts of the ObReco-2. Indeed, at the end of each task, the experimenter required them to remove the Cardboard in order to provide the next link. Here all tasks were provided in a unique VR experience.

Procedure

The study involved randomized between-subject data collection. Each participant performed two sessions, that lasted about one hour, at a maximum of two days apart. In the first session, a neuropsychological assessment was performed, then participants were randomly assigned to different conditions: the traditional paper and pencil tests (RBMT-III Italian Version) and the experimental one (ObReco-2). During the 360° session all the participants were sitting on a turning chair, to freely explore the virtual environments using an Oculus Quest-2© HMD.

Neuropsychological assessment

The neuropsychological evaluation included the MMSE, the Frontal Assessment Battery (FAB) Italian Version (Appollonio et al., 2005), the Babcock Story Recall Test (BSRT) Italian Version (Spinnler and Tognoni, 1987), the Rey Auditory Learning test (RAVLT) (Carlesimo et al., 1996), the Tower of London (ToL) (Allamanno et al., 1987), Attentive matrices (Spinnler and Tognoni, 1987), test exploring Constructive Apraxia (Arrigoni and de Renzi, 1964), Trail Making Test (TMT) (Amodio et al., 2002) and Raven's progressive matrices (Caffarra et al., 2003). Moreover, participants performed the Picture Recognition sub-test included in the RBMT-III Italian Version (Beschin and Urbano, 2013). The Picture Recognition is a sub-test of the RBMT-III. It is divided into two parts: the encoding and the recognition phases. During the encoding, the patient is asked to see a set of 15 pictures representing common animate and inanimate objects (e.g., a clock, a chicken) and to recognize and name each one of them. In the recognition phase, the participant is asked to observe a set of 30 pictures including target items (i.e., the 15 pictures presented in the Encoding Phase) and distractors (i.e., 15 pictures not included in the Encoding Phase): for each of these, the patient is asked to answer yes if the picture was presented previously or no if it was not. During the Recognition task, several measures are collected: the HR (the proportion of yes responses to old items) the False Alarm Rate (the proportion of yes responses to distractors), and the False Alarm Unknown (the proportion of yes responses to unknown distractors) (Snodgrass and Corwin, 1988). The raw score obtained in the sub-test is the number of pictures correctly recognized. Moreover, before the Recognition Phase, we included a Free Recall task, in which the patient was required to recall every object he/she could from those presented

TABLE 3 Qualitative usability results of thinking aloud protocol.

Task	Problem	Solution	N.S.
Use of cardboard			
Wear cardboard	None	None	–
Remove cardboard	Sense of annoyance/sense of falling after removing the cardboard	Encourage the patient to keep his/her eyes open to avoid falling	1
Instructions			
Listening	None	None	–
Comprehension	None	None	–
Familiarization			
Listening	None	None	–
Comprehension	None	None	-
Execution	Blurry image	Improve the quality of VR video	1
Encoding			
Listening	None	None	–
Comprehension	None	None	–
Execution	Difficulty to explore the environment in an appropriate order	Improve instructions' clarity	5
		Encourage to listen carefully the instructions	
	Unclear image	Improve the quality of images	4
	Name all the objects	Improve instructions' clarity	1
	Difficulty to find the initial object labeled	Improve the instructions	4
	Nausea	Provide slower execution of the exercise	1
Recognition			
Listening	None	None	–
Comprehension	None	None	–
Execution	Dizziness	Provide slower execution of the exercise	1
	Recognizes many distractors caused by blurry image	Improve the quality of images	1

in the encoding phase. The raw score is defined by the number of objects correctly reported.

Data analysis

All the analyses were performed using Jamovi Software (The jamovi project, 2021). After having collected all the data in a Windows Excel sheet, we computed different indexes for both the RBMT-III and ObReco-2. In particular, for the recognition tasks, we computed three different scores: the HR, (the proportion of yes responses to targets), the False Alarm Rate (the proporton of yes responses to distractors), and the False Alarm "unknown" (the proportion of yes responses to unknown distractors, i.e., objects that were not included in the encoding phase). Then, we performed Mann–Whitney U tests to compare the free recall and recognition scores in both RBMT-III and 360° modalities, investigating the statistically significant differences in the two performances. We also performed correlation analyses to investigate relationships between neuropsychological examinations and memory indices from RBMT-III and ObReco-2. At last, we performed descriptive analyses of the usability questionnaires and then we compared (using Mann–Whitney U) usability scores of study 1, in which participants used cardboard, and study 2 where otherwise they used an Oculus Quest-2.

Results

Usability

Starting from quantitative data, the mean score of the SUS is 74 (SD = 14.7). According to Bangor et al. (2009) this score indicates that ObReco-2 is placed in a *Good Level* of usability as shown in **Figure 4**.

The results of the STAM scale reveal that our sample has a positive attitude toward technology (M = 6.81/10/10; SD = 2.98), has good control/access to technological devices (M = 7.39/10; SD = 1.66), has a medium level of technology-related anxiety (M = 6/10; SD = 2.81), and considers themselves in good health conditions (M = 7.62/10; SD = 1.31). As shown by the ITC-SOPI sub-scale investigating negative effects, all subjects reported minimal side effects (M = 1.90; SD = 1.79) indicating that the use of ObReco-2 did not determine dizziness and cybersickness. The descriptive of UX measures are shown in **Table 5**. Considering qualitative results of the Thinking Aloud Protocol, a limited number of patients referred to similar problems to those observed in the cardboard's group: they mentioned blurry images and difficulty to identify where the labels are immediately when the task started. At last, we compared the usability scores of study 1 (Cardboard) and study 2 (Oculus Quest-2). The results of the independent *t*-test reveal non-statistically significant differences suggesting that both cardboard and Oculus Quest-2 are easy-to-use technologies.

Clinical efficacy

The descriptives of the accuracy on free recall and recognition tasks performances of two groups are presented in **Table 6**. The results indicate that for the free recall tasks, participants performed better after ObReco-2 than RBMT-III in terms of the number of targets correctly recalled although the difference is not statistically significant (U = 39.0, p = 0.416). Concerning the recognition indexes, participants recognized more objects after the standard presentation compared to the 360° one, and the observed difference is statistically significant (U = 21.5, p = 0.029). We also performed a statistical analysis to investigate correlations between neuropsychological

TABLE 4 Demographic data and mini-mental state examination scores.

Descriptives						
	Years_PP	Education_PP	MMSE_PP	Years_VR	Education_VR	MMSE_VR
Mean	69.7	14.6	27.6	68.2	12.0	27.2
Standard deviation	7.63	3.84	1.79	5.45	4.45	1.74
Min	57	8	25.5	59	5	24.7
Max	81	18	30.0	75	18	30.0

PP, paper and pencil group; VR, virtual reality group.

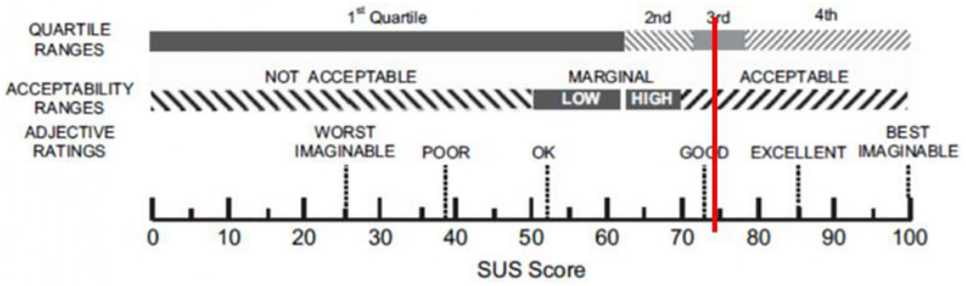


FIGURE 4
Graphical representation of the interpretation of system usability scale (SUS). The vertical line shows the position of the SUS mean score (74) obtained by the Oculus Quest according to the rating comparison scale provided by Bangor et al. (2009).

TABLE 5 Desciptives of the user experience (UX) measures of the study 2.

	SUS	STAM-a	STAM-c	STAM-anx	STAM-h	ITC-sp	ITC-e	ITC-ev	ITC-ne
Mean	74/100	6.81/10	7.39/10	6/10	7.62/10	3.43/5	3.87/5	4.12/5	1.90/5
Standard deviation	14.73	2.98	1.66	2.81	1.32	0.86	0.78	0.71	1.79
Min	55.00	1.00	3.5	2.00	5.60	1.80	2.30	3.00	1.00
Max	90.00	10.00	9.50	10.00	9.20	4.60	4.80	5.00	6.80

For each measure there are mean and the maximum available score, standard deviation and minimum (min) and maximum (max) score reported by participants.
SUS, System Usability Scale; CSQ; STAM-a, attitude through technologies subscale; STAM-c, Senior Technology Acceptance Model perception of control subscale; STAM-anx, Senior Technology Acceptance Model anxiety related to technologies subscale; STAM-h, Senior Technology Acceptance Mode health conditions subscale; ITC-sp, Independent Television Commission Sense of Presence Inventory-Sense of Physical Space subscale; ITC-e, Engagement subscale; ITC-ev, Ecological Validity subscale; ITC-ne, Negative Effects subscale.

examinations and memory indices from RBMT and ObReco-2. In the control group, results show a statistically significant correlation between FAB and RBMT Recognition (HR) ($r = 0.687$, $p = 0.028$). None of the other neuropsychological tests correlates with RBMT. On the other hand, in the experimental group ObReco-2 scores correlate with AM ($r = 0.642$, $p = 0.045$) and delayed RAVLT ($r = 0.645$, $p = 0.044$).

Discussion

The ongoing scientific debate about the ecological validity of classical assessment encourages the implementation of VR in neuropsychological assessment (Chaytor and Schmitter-Edgecombe, 2003; Parsons, 2015; Neguț et al., 2016). Based on this rationale, we aimed to design an assessment tool that used naturalistic and life-like situations; we decided to develop an application using 360°contents which allow a more ecological performance rather than computer-generated VR (Serino and Repetto, 2018). The results are promising: patients were satisfied with the application and they expressed interest in trying a new assessment methodology. They were fascinated by the exploration of a virtual environment, and they reported enjoyment in performing exercises in this innovative way. Furthermore, results revealed minimal negative effects while wearing the cardboard. Only a small number of them experienced dizziness or sickness as a possible collateral effect. However, considering the experience with the cardboard, it was limited: patients could not interact directly with the environment due to technical restrictions. Every phase of the task required the experimenter's intervention and thus the experience of the users was not continuative. To overcome these limitations, we designed and ameliorated this task by employing an advanced technology: the Oculus Quest-2. Results of the

TABLE 6 The table shows the descriptives of the accuracy obtained by the participants in the free recall tasks (FR) and the correct objects identified in the recognition task hit rate (HR) in the standard Rivermead Behavioural Memory Test (RBMT) and virtual reality (VR) (ObReco) conditions. The last column indicates the false alarms (FAR) i.e., the yes responses to the wrong items.

Descriptives

	RBMT_FR	OBRECO_FR	RBMT_HR	OBRECO_HR	OBRECO_FAR
Mean	5.30	6.20	11.1	9.20	1.30
Standard deviation	2.45	2.10	1.52	1.81	0.675

usability scales revealed that ObReco-2 is *Acceptable* and has a *Good Level* of usability (Bangor et al., 2009). It means that the product was judged "goodness" by users. Positive outcomes came also from the results of the STAM scale (Chen and Lou, 2020), confirming users' acceptance and usage. It means that the four key factors identified by the model, *performance expectation*, *effort expectation*, *anxiety related to technology*, *and facilitating conditions* (health condition) predict the intent to use the proposed tool. Technology acceptance is the perception of attitudes and behavioral intent to use technology, and it is a major predictor of technology adoption and usage. In support of this, the Negative Effects scale indicates minimal dizziness and cybersickness. This result could be explained by the minimal movement required in the VR environment and the limited duration of the VR exposure (about 10 min). Nevertheless, some technical problems were reported by the TAP (Lewis, 1982) in both usability studies, possible solutions for the main described issues could be improving the clarity of the instructions, adding a more specific training phase, improving the quality of images, and suggesting a slower execution of the exercise. Although the two study groups were different and we couldn't compare the two devices, the UX measures of both, cardboard and Oculus, do not seem to differ. On one hand, the lack of interactivity of the cardboard could have been experienced as an advantage instead of a limit for a sample of old people who don't have proper skills with technologies. In this way, cardboard may be managed more simply and quickly. On the other hand, the experience provided by the Oculus Quest-2 in terms of immersivity, sense of presence, and engagement guarantee better immersive quality and better VR experience which explain the high scores of acceptability for this device.

A further purpose of this exploratory study was to test the efficacy of 360° technology in neuropsychological memory assessment. Considering previous results from literature (Serino et al., 2017; Realdon et al., 2019) we expected to find some correlations between memory performances in the standard sub-test of the RBMT-III and the ObReco-2. The results indicated that participants obtained higher scores on the free recall tasks in the virtual conditions, showing a better trend performance after the ObReco-2, although this difference is not statistically significant. This trend could be explained by other factors including engagement and interaction provided by the VR experience. In fact, in agreement with previous studies (Robertson et al., 2016; Makowski et al., 2017), a higher level of immersivity and realism leads to better memory encoding. The photorealism of 360° environments may have elicited a visual memory encoding similar to that seen in everyday life, resulting in the greater visual encoding of stimuli, easier recall of items, and increased ecological validity of the evaluation technique. Overall, these results are consistent with those found by Pieri et al., which used a minor number of target objects to be remembered (Pieri et al., 2021). For what concerns the recognition performance, the pattern of results is inverted: participants showed high levels of accuracy in both conditions but performed significantly better in the RBMT-III condition. These results could be explained by analyzing the participants' qualitative reports. They described difficulties in recognizing objects during the encoding phase in the virtual task, due to the low quality of the video. Future studies could introduce a preliminary naming test to verify this condition. Moreover, while in the RBMT-III participants had to encode one object per time, in the virtual task all the target objects were shown in the environment at the same time. This could have prevented them to focus their attention singularly on each object, although this condition is the most similar to real-life situations. This complexity reflects the daily routine in which ecological patterns require actively exploring the space to discriminate the target items from the distractors. This may have allowed a slightly more sensitive and ecological assessment of recognition memory when compared to the RBMT-III condition.

Limitation and conclusion

The present work is not exempt from limitations. First, the sample is restricted in its size and representativity. We primarily focused on the features of the technology, but further studies must include a larger sample size with different demographic characteristics. The second gap regards the technological equipment, currently, the 360° devices market offers much higher-quality omnidirectional cameras (e.g., Insta360 Pro 2©) which can provide a higher-quality of images and a higher ecological value to the obtained measures. Then, another limitation refers to the difference between samples of study 1 and study 2 in terms of MMSE scores.

Even with its limitations, these findings show the feasibility of 360°-VR assessment, thus encouraging the implementation of this technology in the development of ecological tests for memory evaluation. Based on these assumptions, future studies are needed to develop and validate standardized applications for the assessment of different cognitive domains but also different memories, for example, semantic or autobiographical. Further works are also required to clarify which advantages and disadvantages characterize VR, to improve the design of 360° experiences, and to investigate cognitive assessment using the innovative proposed tool in different populations.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethical Commitee-IRCCS Istituto Auxologico Italiano. The patients/participants provided their written informed consent to participate in this study.

Author contributions

FBr designed the protocol and write the first draft. CS-B and VM participated in the design and contributed to the data collection. LG developed the environments. PC, EP, and FBo analyzed the data. MC and MS-B provided the required revisions. GR supervised and contributed to the reviewed version of the manuscript. PC, MS-B, and EP have supervised the study. All authors have approved the final version of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Ministry of Education project "A cross-platform application for the ecological and remote assessment of memory impairment in elderly: ECO-MEMORY" (FISR2020IP_04509) and partially supported by the Italian Ministry of Health (POSTECH: 39C801_2018).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Aida, J., Chau, B., and Dunn, J. (2018). Immersive virtual reality in traumatic brain injury rehabilitation: a literature review. *NeuroRehabilitation* 42, 441–448. doi: 10.3233/NRE-172361

Alashram, A. R., Annino, G., Padua, E., Romagnoli, C., and Mercuri, N. B. (2019). Cognitive rehabilitation post traumatic brain injury: a systematic review for emerging use of virtual reality technology. *J. Clin. Neurosci.* 66, 209–219. doi: 10.1016/j.jocn.2019.04.026

Allain, P., Foloppe, D. A., Besnard, J., Yamaguchi, T., Etcharry-Bouyx, F., Le Gall, D., et al. (2014). Detecting everyday action deficits in Alzheimer's disease using a nonimmersive virtual reality kitchen. *J. Int. Neuropsychol. Soc.* 20, 468–477. doi: 10.1017/S1355617714000344

Allamanno, N., della Sala, S., Laiacona, M., Pasetti, C., and Spinnler, H. (1987). Problem solving ability in aging and dementia: normative data on a non-verbal test. *Ital. J. Neurol. Sci.* 8, 111–119. doi: 10.1007/BF02337583

Amodio, P., Wenin, H., del Piccolo, F., Mapelli, D., Montagnese, S., Pellegrini, A., et al. (2002). Variability of trail making test, symbol digit test and line trait test in normal people. A normative study taking into account age-dependent decline and sociobiological variables. *Aging Clin. Exp. Res.* 14, 117–131. doi: 10.1007/BF03324425

Appollonio, I., Leone, M., Isella, V., Piamarta, F., Consoli, T., Villa, M. L., et al. (2005). The frontal assessment battery (FAB): normative values in an Italian population sample. *Neurol. Sci.* 26, 108–116. doi: 10.1007/s10072-005-0443-4

Arrigoni, G., and de Renzi, E. (1964). Constructional apraxia and hemispheric locus of lesion. *Cortex* 1, 170–197. doi: 10.1016/S0010-9452(64)80020-4

Bangor, A., Kortum, P., and Miller, J. (2009). Determining what individual SUS scores mean: adding an adjective rating scale. *J. Usabil. Stud.* 4, 114–123.

Beschin, N., and Urbano, T. T. B. (2013). *Rivermead Behavioural Memory Test, Adattamento Italiano*, 3rd Edn. Firenze: Giunti Psychometrics.

Bohil, C. J., Alicea, B., and Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nat. Rev. Neurosci.* 12, 752–762. doi: 10.1038/nrn3122

Brooke, J. (2020). "SUS: a "Quick and Dirty" usability scale," in *Usability Evaluation In Industry*, eds P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester (London: CRC Press), 207–212. doi: 10.1201/9781498710411-35

Bruni, F., Mancuso, V., Stramba-Badiale, C., Pieri, L., Riva, G., Cipresso, P., et al. (2022). 360° immersive photos and videos, an ecological approach to memory assessment: the ObReco-2. *Annu. Rev. Cyberther. Telemed.*

Caffarra, P., Vezzadini, G., Zonato, F., Copelli, S., and Venneri, A. (2003). A normative study of a shorter version of Raven's progressive matrices 1938. *Neurol. Sci.* 24, 336–339. doi: 10.1007/s10072-003-0185-0

Carlesimo, G. A., Caltagirone, C., and Gainotti, G. (1996). The mental deterioration battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment. *Eur. Neurol.* 36, 378–384. doi: 10.1159/000117297

Chaytor, N., and Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. *Neuropsychol. Rev.* 13, 181–197. doi: 10.1023/B:NERV.0000009483.91468.fb

Chen, K., and Lou, V. W. Q. (2020). Measuring senior technology acceptance: development of a Brief, 14-Item Scale. *Innov. Aging* 4, 1–12. doi: 10.1093/geroni/igaa016

Laver, K. E., Lange, B., George, S., Deutsch, J. E., Saposnik, G., and Crotty, M. (2017). Virtual reality for stroke rehabilitation. *Cochrane Database Syst. Rev.* 11:CD008349. doi: 10.1002/14651858.CD008349.pub4

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: the ITC-sense of presence inventory. *Presence* 10, 282–297. doi: 10.1162/105474601300343612

Lewis, C. H. (1982). *Using the "Thinking Aloud" Method in Cognitive Interface Design (Technical Report)*. Endicott, NY: IBM.

Magni, E., Binetti, G., Bianchetti, A., Rozzini, R., and Trabucchi, M. (1996). Mini-mental state examination: a normative study in Italian elderly population. *Eur. J. Neurol.* 3, 198–202.

Makowski, D., Sperduti, M., Nicolas, S., and Piolino, P. (2017). "Being there" and remembering it: presence improves memory encoding. *Conscious. Cogn.* 53, 194–202. doi: 10.1016/j.concog.2017.06.015

Matheis, R. J., Schultheis, M. T., Tiersky, L. A., DeLuca, J., Millis, S. R., and Rizzo, A. (2007). Is learning and memory different in a virtual environment? *Clin. Neuropsychol.* 21, 146–161. doi: 10.1080/13854040601100668

Measso, G., Cavarzeran, F., Zappalà, G., Lebowitz, B. D., Crook, T. H., and Pirozzolo, F. J. (1993). The mini-mental state examination: normative study of an Italian random sample. *Dev. Neuropsychol.* 9, 77–85. doi: 10.1080/87565649109540545

Mondini, S., Mapelli, D., and Arcara, G. (2016). *Semeiotica e Diagnosi Neuropsicologica: Metodologia per la Valutazione*. Roma: Carocci.

Neguţ, A., Matu, S., Sava, F. A., and David, D. (2016). Virtual reality measures in neuropsychological assessment: a meta-analytic review. *Clin. Neuropsychol.* 30, 165–184. doi: 10.1080/13854046.2016.1144793

Ouellet, É, Boller, B., Corriveau-Lecavalier, N., Cloutier, S., and Belleville, S. (2018). The virtual shop: a new immersive virtual reality environment and scenario for the assessment of everyday memory. *J. Neurosci. Methods* 303, 126–135. doi: 10.1016/j.jneumeth.2018.03.010

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Front. Hum. Neurosci.* 9:660. doi: 10.3389/fnhum.2015.00660

Parsons, T. D., Carlew, A. R., Magtoto, J., and Stonecipher, K. (2017). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychol. Rehabil.* 27, 777–807. doi: 10.1080/09602011.2015.1109524

Pieri, L., Serino, S., Cipresso, P., Mancuso, V., Riva, G., and Pedroli, E. (2021). The ObReco-360°: a new ecological tool to memory assessment using 360° immersive technology. *Virtual Reality* 26, 639–648. doi: 10.1007/s10055-021-00526-1

Realdon, O., Serino, S., Savazzi, F., Rossetto, F., Cipresso, P., Parsons, T. D., et al. (2019). An ecological measure to screen executive functioning in MS: the Picture Interpretation Test (PIT) 360°. *Sci. Rep.* 9:5690. doi: 10.1038/s41598-019-42201-1

Riva, G., and Mantovani, F. (2014). "Extending the self through the tools and the others: a general framework for presence and social presence in mediated interactions," in *Interacting with Presence: HCI and the Sense of Presence in Computer-Mediated Environments*, eds G. Riva, J. Waterworth, and D. Murray (Warsaw: De Gruyter Open Poland), 9–31. doi: 10.2478/9783110409697.1

Riva, G., Wiederhold, B. K., and Mantovani, F. (2019). Neuroscience of virtual reality: from virtual exposure to embodied medicine. *Cyberpsychol. Behav. Soc. Netw.* 22, 82–96. doi: 10.1089/cyber.2017.29099.gri

Rizzo, A. S., and Koenig, S. T. (2017). Is clinical virtual reality ready for primetime? *Neuropsychology* 31, 877–899. doi: 10.1037/neu0000405

Robertson, C. E., Hermann, K. L., Mynick, A., Kravitz, D. J., and Kanwisher, N. (2016). Neural representations integrate the current field of view with the remembered 360° Panorama in scene-selective cortex. *Curr. Biol.* 26, 2463–2468. doi: 10.1016/j.cub.2016.07.002

Saposnik, G., and Levin, M. (2011). Virtual reality in stroke rehabilitation: a meta-analysis and implications for clinicians. *Stroke* 42, 1380–1386. doi: 10.1161/STROKEAHA.110.605451

Sbordone, R. J. (1996). "Ecological validity: some critical issues for the neuropsychologist," in *Ecological Validity of Neuropsychological Testing*, eds R. J. Sbordone and C. J. Long (Delray Beach, FL: St Lucie Press, Inc), 15–41.

Serino, S., and Repetto, C. (2018). New trends in episodic memory assessment: immersive 360° ecological videos. *Front. Psychol.* 9:1878. doi: 10.3389/fpsyg.2018.01878

Serino, S., Pedroli, E., Tuena, C., De Leo, G., Stramba-Badiale, M., Goulene, K., et al. (2017). A novel virtual reality-based training protocol for the enhancement of the "mental frame syncing" in individuals with Alzheimer's disease: a development-of-concept trial. *Front. Aging Neurosci.* 9:240. doi: 10.3389/fnagi.2017.00240

Snodgrass, J. G., and Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J. Exp. Psychol. Gen.* 117, 34–50. doi: 10.1037/0096-3445.117.1.34

Spinnler, H., and Tognoni, G. (1987). *Standardizzazione e Taratura Italiana Di Test Neuropsicologici: Gruppo Italiano per Lo Studio Neuropsicologico Dell'invecchiamento*. Milano: Masson Italia Periodici.

The jamovi project (2021). *jamovi. (Version 1.6) [Computer Software]*. Available online at: https://www.jamovi.org (accessed January 19, 2022).

Tuena, C., Pedroli, E., Trimarchi, P. D., Gallucci, A., Chiappini, M., Goulene, K., et al. (2020). Usability issues of clinical and research applications of virtual reality in older people: a systematic review. *Front. Hum. Neurosci.* 14:93. doi: 10.3389/fnhum.2020.00093

Ventura, S., Brivio, E., Riva, G., and Baños, R. M. (2019). Immersive versus non-immersive experience: exploring the feasibility of memory assessment through 360 ° technology. *Front. Psychol.* 10:2509. doi: 10.3389/fpsyg.2019.02509

Wilson, B. A., Greenfield, E., Clare, L., Baddeley, A., Cockburn, J., Watson, P., et al. (2008). *The Rivermead Behavioural Memory Test*, 3rd Edn. London: Pearson Education.

Wilson, B., Cockburn, J., Baddeley, A., and Hiorns, R. (1989). The development and validation of a test battery for detecting and monitoring everyday memory problems. *J. Clin. Exp. Neuropsychol.* 11, 855–870. doi: 10.1080/01688638908400940

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 11 August 2022
DOI 10.3389/fnagi.2022.923673

OPEN ACCESS

EDITED BY
Wai Haung Yu,
University of Toronto, Canada

REVIEWED BY
Owen T. Carmichael,
Pennington Biomedical Research Center, United States
Martin Dyrba,
Helmholtz Association of German Research Centers (HZ), Germany

*CORRESPONDENCE
Jia Guo
jg3400@columbia.edu

†These authors have contributed equally to this work
‡Membership of the Alzheimer's Disease Neuroimaging Initiative

SPECIALTY SECTION
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 19 April 2022
ACCEPTED 18 July 2022
PUBLISHED 11 August 2022

CITATION
Liu C, Zhu N, Sun H, Zhang J, Feng X, Gjerswold-Selleck S, Sikka D, Zhu X, Liu X, Nuriel T, Wei H-J, Wu C-C, Vaughan JT, Laine AF, Provenzano FA, Small SA and Guo J (2022) Deep learning of MRI contrast enhancement for mapping cerebral blood volume from single-modal non-contrast scans of aging and Alzheimer's disease brains.
Front. Aging Neurosci. 14:923673.
doi: 10.3389/fnagi.2022.923673

Deep learning of MRI contrast enhancement for mapping cerebral blood volume from single-modal non-contrast scans of aging and Alzheimer's disease brains

Chen Liu[1†], Nanyan Zhu[2†], Haoran Sun[3], Junhao Zhang[3], Xinyang Feng[3], Sabrina Gjerswold-Selleck[3], Dipika Sikka[3], Xuemin Zhu[4], Xueqing Liu[3], Tal Nuriel[5], Hong-Jian Wei[5], Cheng-Chia Wu[5], J. Thomas Vaughan[3], Andrew F. Laine[3], Frank A. Provenzano[6], Scott A. Small[6,7,8] and Jia Guo[7,9*] for the Alzheimer's Disease Neuroimaging Initiative[‡]

[1]Department of Electrical Engineering, Columbia University, New York, NY, United States, [2]Department of Biological Sciences, Columbia University, New York, NY, United States, [3]Department of Biomedical Engineering, Columbia University, New York, NY, United States, [4]Department of Pathology and Cell Biology, Columbia University, New York, NY, United States, [5]Department of Radiation Oncology, Columbia University, New York, NY, United States, [6]Department of Neurology, Columbia University, New York, NY, United States, [7]Department of Psychiatry, Columbia University, New York, NY, United States, [8]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY, United States, [9]The Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, United States

While MRI contrast agents such as those based on Gadolinium are needed for high-resolution mapping of brain metabolism, these contrast agents require intravenous administration, and there are rising concerns over their safety and invasiveness. Furthermore, non-contrast MRI scans are more commonly performed than those with contrast agents and are readily available for analysis in public databases such as the Alzheimer's Disease Neuroimaging Initiative (ADNI). In this article, we hypothesize that a deep learning model, trained using quantitative steady-state contrast-enhanced structural MRI datasets, in mice and humans, can generate contrast-equivalent information from a single non-contrast MRI scan. The model was first trained, optimized, and validated in mice, and was then transferred and adapted to humans. We observe that the model can substitute for Gadolinium-based contrast agents in approximating cerebral blood volume, a quantitative representation of brain activity, at sub-millimeter granularity. Furthermore, we validate the use of our deep-learned prediction maps to identify functional abnormalities in the aging brain using locally obtained MRI scans, and in the brain of patients with Alzheimer's disease using publicly available MRI scans from ADNI.

Since it is derived from a commonly-acquired MRI protocol, this framework has the potential for broad clinical utility and can also be applied retrospectively to research scans across a host of neurological/functional diseases.

KEYWORDS

aging, CBV, deep-learning, gadolinium, MRI, Alzheimer's disease

Introduction

Gadolinium-based contrast agents (GBCAs) are intravenously administered with Magnetic Resonance Imaging (MRI), and they are most widely known for their utility in mapping, enhancing, and detecting structural brain lesions such as those found in cancer, infections, or bleeding (Borges et al., 2012; Lohrke et al., 2016; Shen and Duong, 2016). Another utility of GBCAs, far less popular until recent years, is to identify functional abnormalities, such as those that alter brain metabolism in aging, neuropsychiatric disorders, and neurodegenerative diseases (Belliveau, 1991; Lewandowski et al., 2013; Schobel et al., 2013; Khan et al., 2014). For these functional lesions, GBCAs can be used to generate high-resolution cerebral blood volume (CBV) maps, thereby deriving a quantitative correlate of metabolic dysfunction that is otherwise invisible to MRI without contrast.

However, recent studies have begun to call the safety of GBCAs into question (Quattrocchi and van der Molen, 2017; Ramalho et al., 2017; Guo et al., 2018; Dillman and Davenport, 2020). GBCAs contain gadolinium, a heavy metal, and are injected into a vein to improve the visualization of internal organs, blood vessels, and tissues during an MRI. After being administered, GBCAs are mostly cleared out from the body through the kidneys. However, trace amounts of gadolinium may stay in patients' bodies, including the brain, for months to years after receiving GBCAs. GBCAs may also increase the risk of nephrogenic systemic fibrosis, a rare but serious disease, in people with severe kidney failure. Even if these concerns are addressed, GBCA administration requires intravenous access, a requirement that places risks on patients and healthcare practitioners, as well as limiting its application in cases when contraindicated (Guo et al., 2018). Thus, there is a need to find a "GBCA substitute", one that can generate GBCA-equivalent information from a non-contrast MRI scan.

Apart from other alternatives, one solution may lie in the non-contrast MRI scans themselves. The main purpose of GBCAs is to selectively highlight signals from the blood so that the blood vessels can visually stand out from the surrounding brain tissues. However, though not visually apparent, such blood-tissue contrast is also present even in non-contrast MRI scans. The underlying reason comes from magnetic resonance physics. The intensity of a voxel in a non-contrast MRI scan is determined by the physical properties, namely the proton density (PD), the T1-, T2-, and T2*- relaxation time constants, of the corresponding material within that voxel. Importantly, blood and different brain tissues have distinct T1 relaxation time constants. At 3 Tesla, the average T1 relaxation time constants of the white matter, the gray matter, and the blood are 866.9, 1433.2, and 1984.4 ms, respectively (Hasgall et al., 2022). In addition, a previous study has shown that some dark structures, identified mainly as vessels, are frequently misclassified as the cerebrospinal fluid (CSF) in T2/PD MRI (Dugas-Phocion et al., 2004) and further demonstrated that ignoring vessel contrast when handling partial volume effect can also lead to an over-estimation of the CSF variance in the intensity space. Other studies have also shown that vessels appear darker than brain tissues on T2*-weighted gradient echo magnetic resonance (GRE) images due to shorter T2* relaxation (Small et al., 2000), and this contrast between blood vessels and brain tissues has been named the susceptibility vessel sign (SVS) (Flacke et al., 2000; Rovira et al., 2004). Hence, non-contrast MRI scans are theoretically able to yield patterns of voxel intensities to distinguish blood vessels from surrounding tissues, though such differences are too subtle to reliably detect and quantify with previous analytical or qualitative methods. Nevertheless, according to magnetic resonance physics, at least a fraction of the GBCA-contrast information due to blood vessels is present and partially encoded, in non-contrast structural MRI scans through a non-linear function.

Deep learning, a subset of machine learning, is an established method for approximating non-linear functions using a data-driven approach. A deep learning model should, therefore, be able to learn how to optimally extract key features at a voxel level, by inspecting MRI scans where GBCAs were administered along with their non-contrast counterparts. As such, a growing number of recent studies have begun validating this assumption (Kleesiek et al., 2019; Liu et al., 2019; Li et al., 2021). Among these, one study managed to use deep learning to reduce the GBCA dose (Gong et al., 2018), but not to completely substitute for it. Other studies succeeded in obviating the need for GBCA (Kleesiek et al., 2019; Liu et al., 2019; Li et al., 2021) but these deep learning models require the acquisition of an array of multiple MRI sequences, some of which are not widely or clinically available. Among these studies that succeeded in obviating the need for GBCA, more

modern deep learning-based methods are especially well suited for this task and demonstrated favorable performance: for the BayesUNet method (Kleesiek et al., 2019) with a comprehensive multiparametric MRI protocol including pre-contrast T1-weighted, T2-weighted, T2-weighted fluid-attenuated inversion recovery (FLAIR), diffusion-weighted imaging (DWI) and susceptibility-weighted imaging (SWI) sequences as the input for predicting synthetic post-contrast T1-weighted sequence, analysis of the whole brain showed a peak signal-to-noise ratio (PSNR) of 22.967 $\pm$ 1.162 and a structural similarity index (SSIM) of 0.872 $\pm$ 0.031; for the CGAN method (Preetha et al., 2021) with pre-contrast T1-weighted, T2-weighted and FLAIR sequences as the input, the model reached a median SSIM of 0.818 (95% CI 0.817 - 0.820); and the MMgSN-Net method (Li et al., 2021) with pre-contrast T1-weighted and T2-weighted sequences as the input achieved top-ranked scores in averaged PSNR of 33.17 $\pm$ 2.14 and SSIM of 0.887 $\pm$ 0.042.

With these issues in mind, we hypothesized that a deep learning model could extract GBCA-equivalent information from a single and commonly-acquired high-resolution MRI scan, by training and optimizing the model using a large and unique GBCA MRI dataset. Previous deep learning studies relied on GBCA datasets generated for radiological purposes, where post-GBCA scans are, by necessity, re-scaled in order to facilitate a radiologist's ability to detect and characterize brain abnormalities. Such re-scaling operations are performed in a case-by-case manner without a universal scaling factor, thus increasing the intersubject variability across a dataset.

Through our previous study in mapping functional brain lesions that localize to specific regions of the hippocampal formation, we have extensively used GBCAs to generate quantitative, high-resolution CBV maps (Small et al., 2011; Pavlopoulos et al., 2013; Schobel et al., 2013; Brickman et al., 2014; Khan et al., 2014; Provenzano et al., 2020). By design, these quantitative maps preserve scaling with respect to the post-GBCA image. While not the original intent, we have accrued a large-scale dataset with reduced inter and intrasubject variability, which we predicted would benefit the training of our model. In parallel to generating a large-scale and quantitative GBCA dataset in humans, we have also accumulated a similar MRI dataset in mice (Moreno et al., 2006; Khan et al., 2014). Again, the original intent was to validate patterns of hippocampal dysfunction observed across disease states; however, because these animal study subjects were siblings with identical genetic backgrounds, this dataset is likely to contain less variability than possible in humans.

In this study, we exploited this distinct cross-species and quantitative GBCA dataset. Beginning with mice to prove the concept, we first designed, optimized, and trained a deep learning model to synthesize GBCA enhancement in the mouse brain from the T2-weighted structural MRI. We further adapted the proposed deep learning model to the human dataset and validated that it can also indicate GBCA enhancement in the human brain from the T1-weighted MRI. The deep learning model will be referred to herein as "DeepContrast". We then utilized this DeepContrast technique to study brain aging and Alzheimer's disease, applying it to both in-house datasets from an aging study, as well as to the publicly available ADNI dataset from Alzheimer's patients and age-matched controls. The studies conducted are outlined in Figure 1. Our results demonstrate that trained deep learning contrast enhancement models can successfully identify and localize brain functional changes that occur through aging and Alzheimer's disease previously only identifiable with GBCA methods.

Materials and methods

As an overview, we conducted 4 sets of studies, as summarized in Figure 1: Healthy Mouse, Healthy Human, Human Aging, and Human AD. In the first two studies, we qualitatively and quantitatively assessed the ability of our proposed DeepContrast model to predict the GBCA enhancement in healthy mice and human brains. In the latter two, we first conducted pilot studies to validate whether the DeepContrast-synthesized CBV maps exhibited the same patterns observed from real CBV data in prior research. Then we performed utility studies to demonstrate the potential use cases. Due to the inherent differences between the Aging and AD studies (presence vs. absence of ground truth, continuous vs. categorical, etc.), the utility studies were designed differently.

In this section, we will describe the data source, data preparation, and detailed analyses performed.

Animal subjects and human participants

Healthy mouse

We used 49 healthy adult C576J/BL male mice (12–14 months old).

Healthy human

We aggregated the healthy human MRI data from previous acquisitions at Columbia University. As we mentioned in the Introduction section, these scans were originally acquired for the purposes of mapping functional lesions. This included 598 participants (16–94 years old) with single acquisitions and another 11 participants with baseline and follow-up acquisitions 14 days apart.

Human aging

For both the pilot and utility studies, we used scans from 177 participants (20–72 years old) that were cognitively

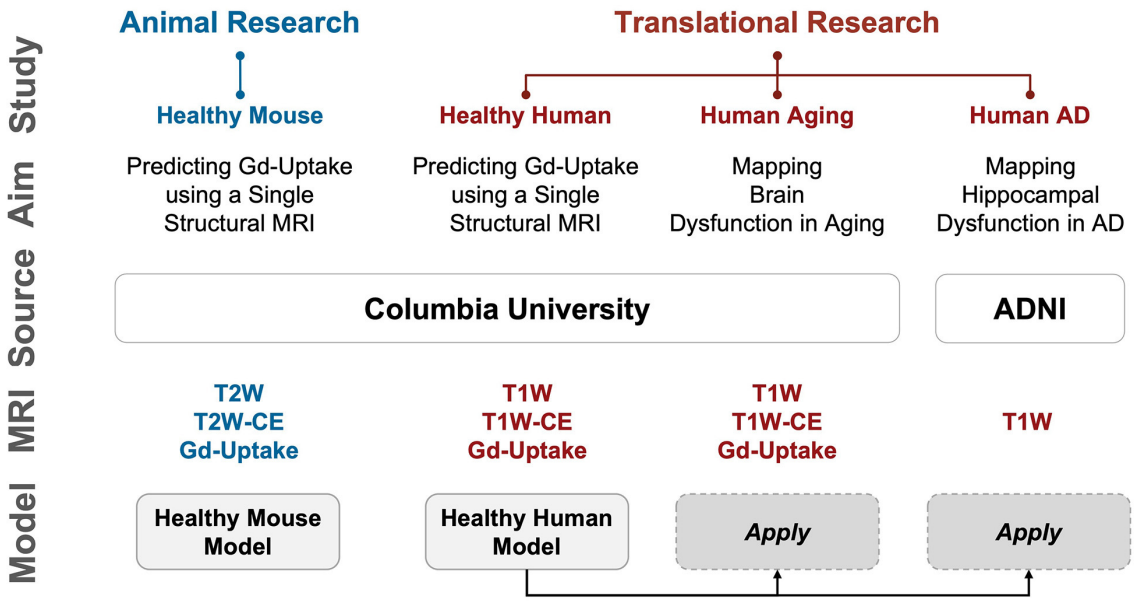


FIGURE 1
Overview of the studies conducted. We first performed proof-of-concept studies in mice to validate our hypothesis that deep learning can extract information equivalent to Gadolinium-based contrast agent (GBCA) contrast enhancement from a single-modal non-contrast MRI scan, and then conducted extensive analyses in humans to scrutinize the capability of this proposed approach. Study: A study conducted; Aim: The purpose of the study; Source: where the imaging data come from; MRI: modality/type of data used in the study; Model: specific DeepContrast model used in the study. AD: Alzheimer's disease; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset; CBV: cerebral blood volume; Gd-Uptake: GBCA contrast uptake maps; T2W: T2-weighted scans; T2W-CE: T2-weighted contrast-enhanced scans; T1W: T1-weighted scans; T1W-CE: T1-weighted contrast-enhanced scans.

normal. These 177 participants were a subset of the 598 healthy participants. The selection criteria will be described in the following 'Preprocessing and partitioning' section.

Human AD

For the pilot study, we selected 50 cognitively normal (CN) and 50 Alzheimer's disease (AD) participants, each with two back-to-back repeated scans from ADNI (Mueller et al., 2005), resulting in a 100-participant (60–90 years old) dataset. For the utility study, we prepared a larger, 2,580-participant (50–100 years old) dataset from ADNI, with 1290 CN and 1290 AD participants.

Image acquisition protocols

Healthy mouse

We used CBV-fMRI to image male Wildtype (WT) mice used in the healthy mouse study, with the imaging protocol as previously described (Moreno et al., 2006). A Bruker BioSpec 94/30 (field strength, 9.4 T; bore size, 30 cm) horizontal small animal MRI scanner equipped with CryoProbe and software ParaVision 6.0.1 (Bruker BioSpin, Billerica, MA, USA) and a 23-mm 1H circularly polarized transmit/receive capable mouse head volume coil were used for the imaging. Mice were anesthetized using medical air and isoflurane (3% volume for induction, 1.1–1.5% for maintenance at 1 liter/min air flow, *via* a nose cone). A flowing water heating pad was used to maintain the body temperature at around 37°C. Sterile eye lubricant was applied before each scan. T2-weighted images were acquired before and 36 min after intraperitoneal injections of the GBCA-based contrast agent Gadodiamide (Omniscan; GE Healthcare, Princeton, NJ, USA) at the dosage of 10 mmol/kg. T2-weighted images were acquired with Refocused Echoes (RARE) sequence (repetition time (TR) = 3,500 ms, effective echo time (TE) = 45 ms, rapid acquisition and relaxation enhancement (RARE) factor = 8, voxel size = 450×76×76 μm).

Healthy human

The images were acquired under a steady-state CBV-fMRI protocol as previously described (Khan et al., 2014). A gradient echo T1-weighted scan (TR = 6.7 ms, TE = 3.1 ms, field of view (FOV) = 240×240×192 mm, voxel size = 0.9×0.9×0.9 mm) was acquired before a pair of un-scaled T1-weighted images (TR = 7 ms, TE = 3 ms, FOV = 240×240×196 mm, voxel size = 0.68×0.68×3 mm), all using a Philips Achieva 3.0-T MRI scanner. The image resolution used results from a systematic exploration of the scan

protocol's parameters. Scans were acquired before and after a bolus injection of a GBCA-based contrast agent (Omniscan, GE Healthcare).

Human aging

The data used in the Human Aging study was a subset of the Healthy Human study, and hence the protocols were identical.

Human AD

The images included in our studies were acquired using a 3D magnetization prepared rapid gradient echo (MP-RAGE) protocol, yielding near-isotropic images (voxel size $\approx 1\times1\times1$ mm). More details can be found in the official documentation of ADNI (Mueller et al., 2005).

Preprocessing and partitioning

Healthy mouse

In total, 49 WT mice were used in this study. Whole brain T2W MRI scans before (T2W) and 35 mins after (T2W-CE) intraperitoneal injection of Gadodiamide were acquired. The Gd-Uptake ground truth was quantified with the standardized delta-R2, which was derived using the same method as discussed before (Moreno et al., 2006), followed by standardization to the dynamic range of [0, 1]. We used 3D PCNN (Chou et al., 2011) with the manual correction to generate brain masks, which we used as training fields over which the model was optimized and performance metrics were calculated. A train-validation-test ratio of 8:1:1 was applied in the Healthy Mouse Model training.

Healthy human

T1-weighted MRI scans were acquired using the protocols as described previously (Brickman et al., 2014; Provenzano et al., 2020), before (T1W) and 4 min after (T1W-CE) a bolus intravenous injection of Gadodiamide. Unlike many other similar studies, during the MRI acquisition for the same session, the receiver gain was intentionally kept constant and the offset was set to zero. As a result, the T1W and T1W-CE scans shared the same scaling and zero shifting, and hence the same voxel intensity between each T1W/T1W-CE pair corresponds to the same relaxation-time property in the magnetic resonance physics context. Each T1W and T1W-CE pair was spatially aligned when provided. For intensity normalization, each T1W scan was compressed to the dynamic range of [0, 1], and the corresponding T1W-CE scan was scaled by the same factor to preserve the voxel intensity correspondence. The Gd-Uptake ground truth was quantified with the steady-state MRI method (Brickman et al., 2014), by subtracting the normalized T1W scans from the respective T1W-CE scans. We generated brain masks using the BET function in FMRIB Software Library (FSL) (Jenkinson et al., 2012), which we used as training fields over which the model was optimized and performance metrics were calculated. We generated tissue label maps using the FAST function in FSL for tissue-of-interest analyses. The train-validation-test split yielded 326 for training, and 93 for validation, while 179 participants were left for the test set.

Human aging

The 177-participant cohort used for the aging study was a subset of the 179 participants in the test set of the Healthy Human Model, where 2 participants were dropped due to low segmentation quality as defined through a failure of processing the FreeSurfer (v6.0.0) Parcellation. After normalization to the dynamic range of [0, 1], the T1W scans were directly treated as inputs to the model to generate the Gd-Predicted maps. Synthesized CBV maps were then generated by applying the same normalization method on the Gd-Predicted maps as we would quantify CBV maps.

Human AD

For the large-scale utility study, we screened T1W MRI scans and excluded all scans except for 3 Tesla MP-RAGE acquisitions (Supplementary Figure S4b top left). After that, we further performed propensity score matching (PSM) to match the age distribution and eventually resulted in a dataset with 1,290 scans of patients with AD and 1,290 scans of age-matched CN volunteers (Supplementary Figure S4b bottom left). A major challenge was that the appearance and anatomy of the scans used in the AD study notably differ from those used to train the DeepContrast Healthy Human Model. They were acquired under the same field strength (i.e., 3 Tesla), but specific scan parameters such as echo time and repetition time are different between the ADNI protocol and the CBV-fMRI protocol. Additionally, the participants in the AD study are generally older (60–90 years old) and half of them harbor Alzheimer's pathology, thus resulting in a potential mismatch in anatomy. We approached these issues by applying (1) affine registration for T1W MRI data and (2) rigid registration to the unbiased MNI152 template on the raw whole brain data and then (3) minimizing the between-cohort appearance difference using a dynamic histogram warping (DHW) algorithm (Cox et al., 1995) as it was demonstrated to be among the best intensity matching methods in medical imaging (Wagenknecht et al., 2000). Specifically, we calculated the mean normalized-brain-region 2048-bin histogram of each cohort derived a bin-to-bin mapping between the cohorts and applied the mapping to each individual scan in the AD study. In step (4), we minimized the anatomical difference by diffeomorphic registration using the Symmetric Normalization (SyN) algorithm (Avants et al., 2009) prior to applying the DeepContrast model. Finally, we

normalized the scans to the dynamic range of [0, 1] and provided them to the model to generate the Gd-Predicted maps. Synthesized CBV maps were then generated by applying the same normalization method as we would quantify synthesized CBV maps and up-sampled to a voxel size of 1×1×1 mm. The data pre-processing is illustrated in Supplementary Figure S4a. The prepared cohort with a total of 2580 T1W scans and 2580 synthesized CBV scans, were randomly assigned to train, validation, and test sets at an 8:1:1 ratio. Randomization was performed on the participant level to prevent data leakage. AD and CN participants were independently randomized to balance the presence of both classes in each set. The data distribution was summarized in the right half of the Supplementary Figure S4b.

For the pilot study, we used a subset containing 50 AD and 50 CN participants each with two back-to-back repeated scans. Compared to the large-scale utility study, the sample size was reduced to accommodate the voxel-based data analysis tool (SPM12) used in the pilot study. Data processing was the same as above.

DeepContrast model implementation

All model variants developed in our studies, as mentioned in Figure 1, shared the common residual attention U-Net (RAU-Net) architecture (Figure 2). Model inputs were the non-contrast MRI scans, while the outputs were the corresponding predicted GBCA contrast (Gd-Predicted). The inputs and outputs were 2D slices of equal dimension since the MRI scans were acquired under 2D protocols. The slice direction was defined as the axis with the lowest spatial resolution, which was axial for Healthy Mouse scans and coronal for Healthy Human scans.

The RAU-Net is an extension of the arguably most popular deep learning architecture in medical imaging, the U-Net (Ronneberger et al., 2015), with the addition of residual blocks (He et al., 2016) and the attention gates (Vaswani et al., 2017; Oktay et al., 2018). As an example of a convolutional neural network (CNN), the U-Net extracts imaging features by utilizing local convolutions along the entire image or volume. The U-Net consists of multiple encoding layers across which the image dimension shrinks whereas the feature dimension increases so that compact high-level abstractions are generated along the process, and the same number of decoding layers to decipher these abstractions into image space information. The add-on residual blocks simplify the entities to be approximated across each layer and, therefore, enables training of deeper networks, while the attention gates learn to differentially enhance or suppress specific regions in the feature maps so that the downstream outcomes are better represented for targeting objective.

Specifically, the encoding and decoding paths consist of the same number of residual convolution blocks that utilize concatenation, attention mechanisms, and skip connections such that layers feed not only into the next layer but into the layer after the next layer. On the encoding path, each residual block is followed by a max-pooling layer, and the last feature map feeds into a bottleneck layer with 3×3 convolution and batch normalization, connecting the deepest layer to the decoding path with several more blocks alternating one un-pooling layer and one residual block. Skip connections concatenate the output of each dense layer in the encoding path with the respective un-pooled feature map of the same size before feeding it as input to the decoding residual block. The output of the last decoding layer is the input for a 1×1 convolution layer that produces the final Gd-Predicted map.

Healthy mouse model

The model (Supplementary Figure S1) used in mouse studies was a 2D RAU-Net that consisted of 5 encoding and decoding layers. The model input was a 2D axial slice of the mouse brain scans. Adam optimizer with a learning rate of 0.001 was used in this study. Our batch size was 3 and the loss function was mean squared error (MSE).

Healthy human model

The model (Supplementary Figure S2) used in the healthy human study and further applied to the Aging and AD studies was a 2D RAU-Net that consisted of 6 encoding and decoding layers. The model input was a 2D coronal slice of the human brain scans. SGD optimizer with an adaptive learning rate handle with a 0.1 initial learning rate was used in this study. Our batch size was 4 and a robust adaptive loss function (Barron, 2019) was utilized. The robust adaptive loss function is a generalization of the Cauchy/Lorentzian, Geman-McClure, Welsch/Leclerc, generalized Charbonnier, Charbonnier/pseudo-Huber/L1-L2, and L2 loss functions. By introducing robustness as a continuous parameter, the robust adaptive loss function allows algorithms built around robust loss minimization to be generalized, which improves performance on basic vision tasks like calculating the intensity mapping function in our case.

Statistical methods

Healthy mouse and healthy human

Prediction vs. ground truth similarity assessment

Peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) (Wang et al., 2004), Pearson correlation coefficient (P.R), and Spearman correlation coefficient (SR) were used to quantify the performance of all the DeepContrast models. PSNR, Pearson correlation coefficient, and Spearman correlation coefficient were evaluated within the brains or

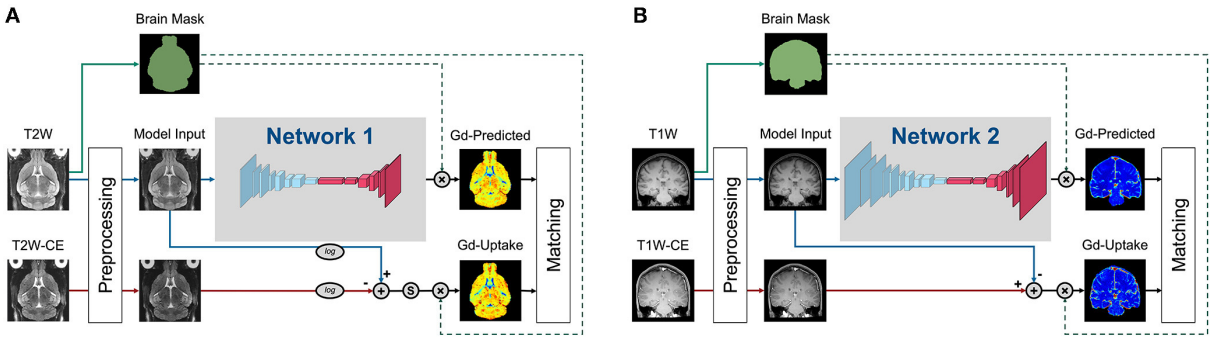


FIGURE 2
Training strategies of the various DeepContrast models implemented. **(A)** This training strategy applied to the Healthy Mouse Model. Preprocessing includes intensity normalization and brain extraction. Ground truth Gd-Uptake was derived using the standardized delta-R2 equation. Note that there is an additional standardization step that maps the dynamic range of the standardized delta-R2 to the range of [0, 1], before the application of the brain mask. The loss function was calculated between the Gd-Uptake and the predicted version only using the voxels within the brain mask region. **(B)** This training strategy applied to the Healthy Human Model. Preprocessing included intensity normalization and brain extraction. Ground truth Gd-Uptake was derived using the steady-state delta-R1 equation. The loss function was calculated between the Gd-Uptake and the predicted version only using the voxels within the brain mask region.

subregions, and SSIM was calculated in the minimum bounding box around the brains or subregions.

Human aging

Pilot study part 1: Voxel-based analysis on the hippocampal circuit

Voxel-based analysis (Figures 4B,C) was performed by first transforming the non-contrast images using a diffeomorphic registration algorithm (Avants et al., 2009) with nearest-neighbor interpolation to an unbiased brain template created from the 177 scans in the Aging study (Avants et al., 2009). The GBCA-predicted maps were generated by the Healthy Human model using the native-space non-contrast T1W scans as the input and were subsequently used to generate synthesized CBV maps by normalization using the mean value among the top 10% brightest voxels within the brain region (representing signal intensity from pure blood). These synthesized CBV maps were then transformed into the template using the same transformation parameters calculated from the registration process and subsequently smoothed using a 3 mm-diameter spherical kernel. Transformed and filtered synthesized CBV maps were analyzed using SPM12 (Ashburner et al., 2014). Data were analyzed with a multiple regression model, including sex as a covariate and age as the regressor. Age-related differences were contrasted using Student's *t*-test. FreeSurfer regional segmentation was then performed on the unbiased template image, and the hippocampal formation mask is generated by binarizing and combining the labels corresponding to the hippocampus and entorhinal cortex. The age-related regression *t*-map was then projected onto the MNI-152 brain template using diffeomorphic transformation with nearest-neighbor interpolation. The result was thresholded at $p < 0.005$ and corrected for multiple comparisons at the cluster level within the hippocampal formation using a Monte-Carlo simulation implemented in AFNI-3dClustSim (Forman et al., 1995; Cox, 1996; Cox et al., 2017) (10,000 iterations) to yield a corrected $p < 0.05$. The final corrected age-related regression t-map was then overlaid onto the MNI-152 template in cross-section using 3DSlicer (Fedorov et al., 2012) and also displayed with composite-with-shading volume rendering over semi-transparent models of the hippocampal formation.

Pilot study part 2: Region of interest analysis on aging-related dentate gyrus region

The 177 native-space synthesized CBV scans were used to conduct the dentate gyrus (DG) region of interest (ROI) analysis. Multiple linear regression with sex as a covariate and age as the regressor was conducted over the bilateral DG, as defined by FreeSurfer parcellation. A scatter plot was drawn (Figure 4D) with each point representing the DG-mean synthesized CBV value after the removal of the sex effect for one participant.

Utility study: Synthesized CBV maps aging effects over the entire cortex

The GBCA-predicted maps were generated in the native space of each participant and were afterward used for CBV quantification together with the experimentally acquired ground truth GBCA-uptake maps using the same whole brain top 10% mean normalization. Similarly, the T1W scans were normalized to generate a comparable counterpart. We used T1W scans for comparison because they were the only input to the DeepContrast model to generate GBCA-predicted maps. The CBV (quantified from Gd-Uptake), synthesized CBV

10.3389/fnagi.2022.923673

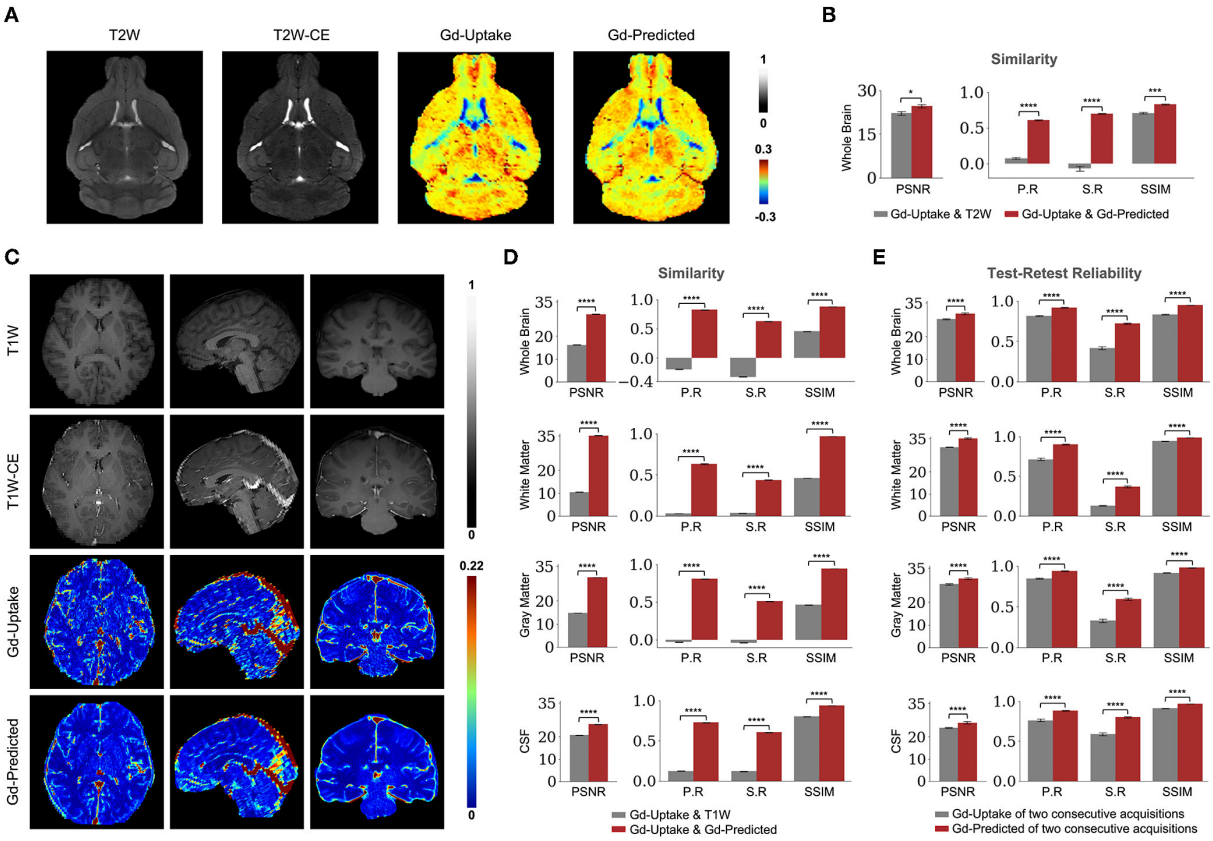


FIGURE 3
Quantitative evaluation of the DeepContrast in the healthy mouse and human brains. **(A)** DeepContrast prediction (Gd-Predicted) highly concords with the ground truth GBCA-uptake map (Gd-Uptake) in the mouse brain. The non-contrast scans and the contrast-enhanced scans are displayed for reference. Color bars indicate the colormap and dynamic range used in the cross-sectional brain images. **(B)** The similarity between the model prediction and the ground truth, evaluated on all 6 scans in the test set using quantitative metrics, where the non-contrast (T2W) scans are used as the performance baseline. **(C)** DeepContrast prediction (Gd-Predicted) highly concords with the ground truth GBCA-uptake map (Gd-Uptake) in the cognitive normal human brain. Color bars indicate the colormap and dynamic range used in the cross-sectional brain images. **(D)** The similarity between the model prediction and the ground truth, evaluated on 179 scans of cognitively normal (CN) participants using quantitative metrics, where non-contrast (T1W) scans are used as the performance baseline. **(E)** DeepContrast shows higher test-retest reliability than the experimentally acquired Gd-Uptake ground truth. For all voxel-based metrics, only the voxels within the brains or subregions are used. SSIM is calculated on the minimum bounding box of the brains or subregions. Asterisks indicate level of statistical significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$). PSNR: peak signal-to-noise ratio; SSIM, structural similarity index; P. R, Pearson correlation coefficient; S.R, Spearman correlation coefficient.

(quantified from Gd-Predicted), and normalized T1W scans were used for age-related regression in the multiple brain regions. Multiple linear regressions with sex as a covariate and age as the regressor were conducted using the mean CBV/synthesized CBV/T1W values extracted from the region across 177 participants, over selected regions (Figure 5) and overall 72 cortical ROIs (Figure 6). The ROIs were parcellated by FreeSurfer over the T1W scans in the native space in order to minimize segmentation errors.

For the ROC analysis, each ROC figure contained 1,000 individual ROC curves. The average ROC was shown as a solid black curve while the SD was shown as the shaded area. All these individual ROC curves were computed using one pair of ground truth (CBV) *t*-score maps and a prediction candidate (synthesized CBV or non-contrast T1W) *t*-score map. Both the ground truth *t*-score map and the prediction candidate *t*-score map were binarized into 2 classes at 1,000 different binarization thresholds evenly distributed between the minimum and the maximum value, yielding 1,000 versions for each. Each individual ROC curve was derived using the regular ROC computation method as described above with one of the 1,000 versions of the ground truth and all 1,000 versions of the prediction candidate. The ROC analysis was performed using Scikit-learn (Pedregosa et al., 2011).

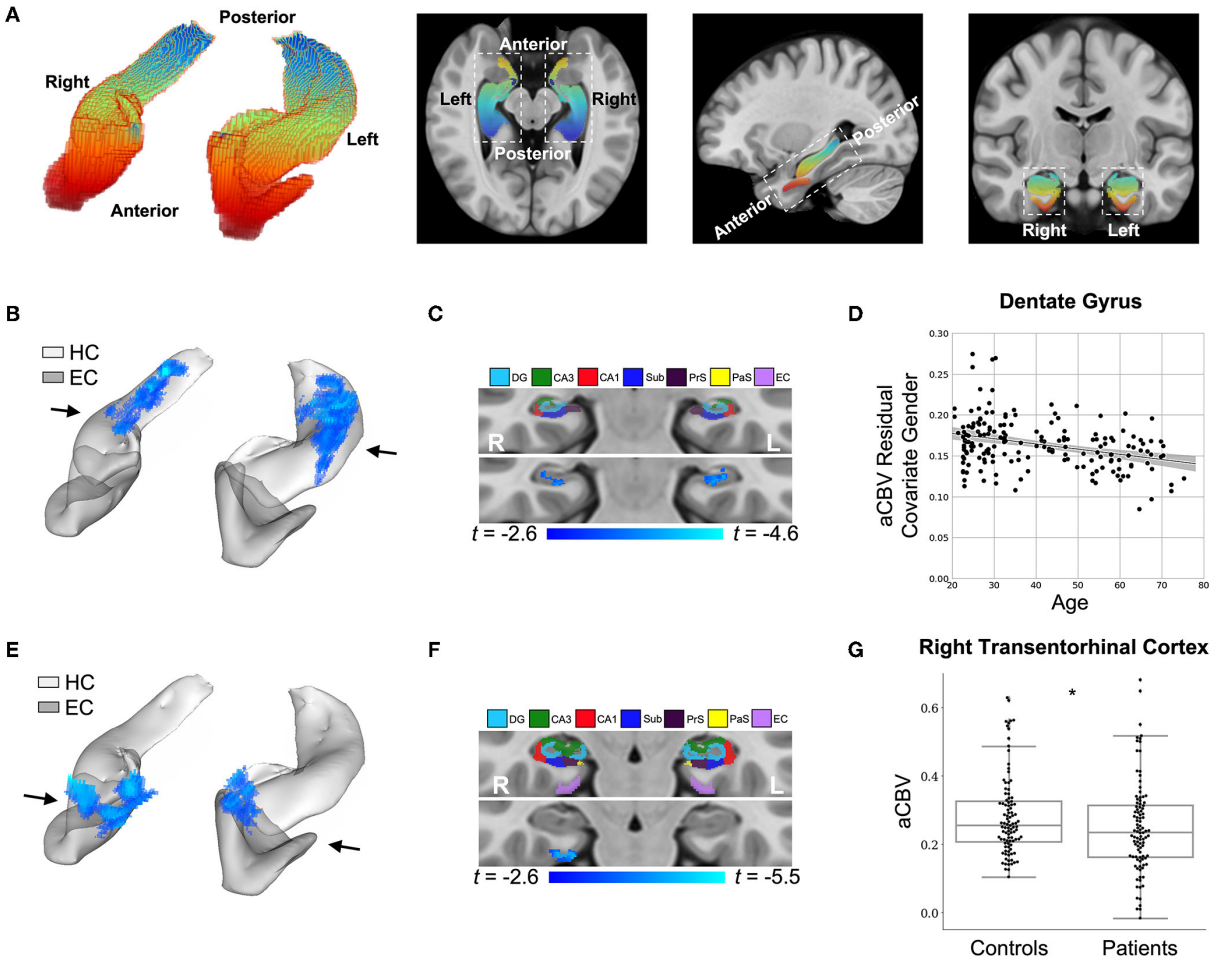


FIGURE 4
DeepContrast maps differential anatomical patterns of dysfunction in the hippocampal formation. **(A)** A three-dimensional rendering of the bilateral hippocampal formation (left panel) consisting of the hippocampus (HC) and the entorhinal cortex (EC) and axial, sagittal, and coronal slices from a group-wise T1-weighted MRI template cutting through the hippocampal formation (right three panels). The hippocampal formation is displayed along the anterior-to-posterior axis. **(B)** A voxel-based analysis of the synthesized CBV maps of 177 individuals ranging from 20 to 72 years of age reveals that the greatest age-related decline occurred in the body of the hippocampal circuit (color-coded by the degree of significance). **(C)** A coronal slice, onto which the hippocampal formation mask is applied, reveals that age-related decline primarily localizes to the dentate gyrus. The voxel-based analysis is conducted using a multiple regression model in SPM12 using sex as a covariate and age as the regressor, and the age-related differences are contrasted using Student's *t* test. Multiple comparisons are corrected for, yielding voxel-wise $p < 0.005$ and cluster-wise $p < 0.05$ (refer to methods). **(D)** A scatter plot shows the association between age and mean synthesized CBV values in the dentate gyrus after the removal of gender effects (β_{age} = −6.36e-4, t_{age} = −4.64, p_{age} = 6.85e-6). The shaded area surrounding the regression line indicates the 95% CI. **(E)** A voxel-based analysis of the synthesized CBV maps of 50 Alzheimer's disease (AD) patients compared with 50 age-matched normal controls, each with two back-to-back scans, reveals AD-related reduction in the entorhinal cortex (color-coded by the degree of significance). **(F)** A coronal slice, onto which the hippocampal formation mask is applied, reveals that AD-related decline localizes primarily to the transentorhinal cortex. The voxel-based analysis is conducted using a multiple regression model in SPM12 using age, sex, and participant identity as covariates and diagnostic class (i.e., cognitive normal vs. dementia) as the regressor and the AD-related difference are contrasted using Student's *t*-test. Multiple comparisons are corrected for, yielding voxel-wise $p < 0.005$ and cluster-wise $p < 0.05$ (refer to methods). **(G)** A box plot showing individual-participant mean synthesized CBV values in the right transentorhinal cortex indicates a significant difference between patients with Alzheimer's disease and healthy controls (two sample *t*-test one-tailed $p = 0.031$). Center line: median; box limits: upper and lower quartiles; whiskers: 1.5× interquartile range; points: outliers. HC: hippocampus; EC: entorhinal cortex; DG: dentate gyrus; CA3: *cornu Ammonis* 3; CA1: *cornu Ammonis* 1; Sub: subiculum; Prs: presubiculum; PaS: parasubiculum.

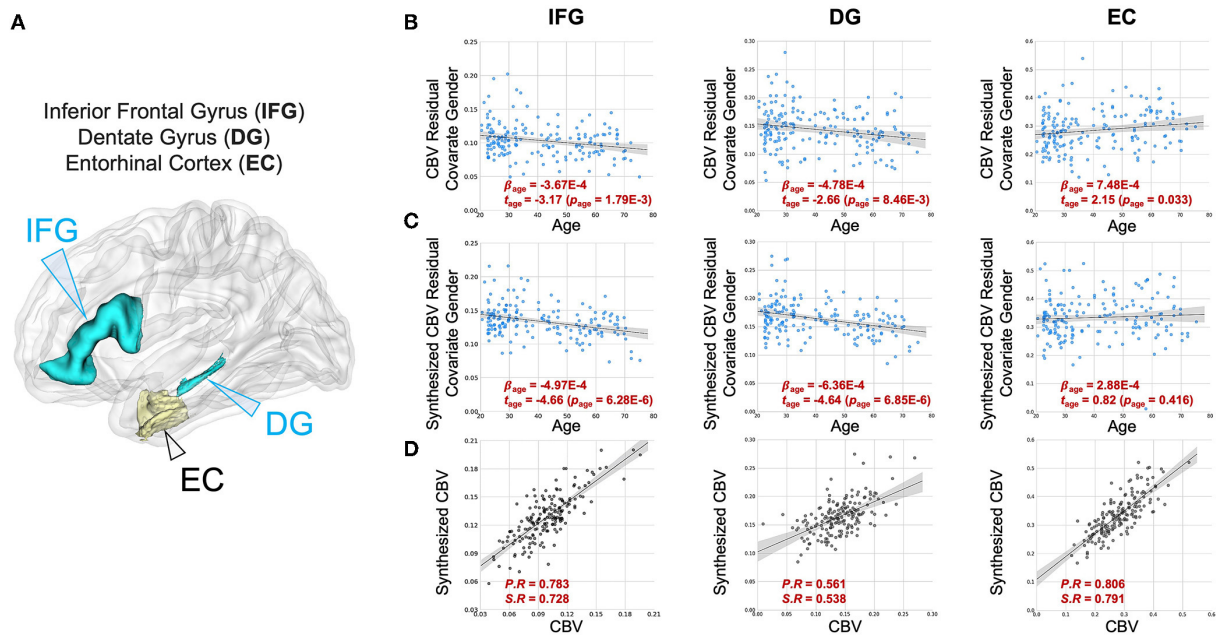


FIGURE 5
DeepContrast maps age-related changes in brain regions vulnerable and resistant to aging. **(A)** A three-dimensional rendering of the inferior frontal gyrus (IFG), dentate gyrus (DG) and entorhinal cortex (EC) overlaid on a group-wise T1-weighted MRI template. **(B)** The age-related regressions of CBV maps over these regions demonstrate the aging-vulnerability of IFG and DG and the aging-resistance of EC. **(C)** The age-related regressions of synthesized CBV maps over the same regions demonstrated the same vulnerability or resistance to aging. **(D)** The scatter plots of the ROI-mean CBV vs. synthesized CBV values of the 177 participants further show the consistency of the two measures.

Human AD

Pilot study part 1: Voxel-based analysis on the hippocampal circuit

Voxel-based analysis (Figures 4E,F) was performed by first transforming the non-contrast images using a diffeomorphic registration algorithm (Avants et al., 2009) with nearest-neighbor interpolation to an unbiased brain template created from the 200 scans (i.e., 50 AD and 50 CN participants each with 2 back-to-back repeated scans) in the pilot study. We then ran these non-contrast scans through the DeepContrast Healthy Human Model to generate synthesized CBV maps, which were subsequently smoothed using a 3 mm-diameter spherical kernel. Unlike in the aging study, the application of DeepContrast was performed after the registration process to help eliminate major anatomical variances, since the deformations present in the diseased population were not previously observed by the model trained on healthy data. GBCA-predicted scans, the direct output of the model, were used to quantify synthesized CBV maps using the same method as described in the Aging study above. These synthesized CBV maps, already co-registered upon creation, were analyzed using SPM12. Data were analyzed with a multiple regression model, including age and sex as covariates and diagnostic class (i.e., cognitive normal vs. dementia) as the regressor. AD-related differences were contrasted using Student's *t*-test. FreeSurfer regional segmentation was then performed on the unbiased template image, and the hippocampal formation mask was generated by binarizing and combining the labels corresponding to the hippocampus and the entorhinal cortex, while an extended hippocampal formation mask was additionally generated to also include the parahippocampal cortex. The AD-related regression t-map was then projected onto the MNI-152 brain template using diffeomorphic transformation with nearest-neighbor interpolation. The result was thresholded at $p < 0.005$ and corrected for multiple comparisons at the cluster level within the extended hippocampal formation using a Monte-Carlo simulation implemented in AFNI-3dClustSim (10,000 iterations) to yield a corrected $p < 0.05$. The final corrected AD-related regression t-map was then overlaid onto the MNI-152 template in cross-section using 3DSlicer and also displayed with composite-with-shading volume rendering over semi-transparent models of the hippocampal formation.

Pilot study part 2: Region of interest analysis on AD-related transentorhinal cortex region

The 200 template-space synthesized CBV scans were used to conduct the right transentorhinal cortex (TEC) ROI analysis. A two-sample *t*-test was conducted over the right TEC, at the boundary between the right entorhinal cortex (EC) and the

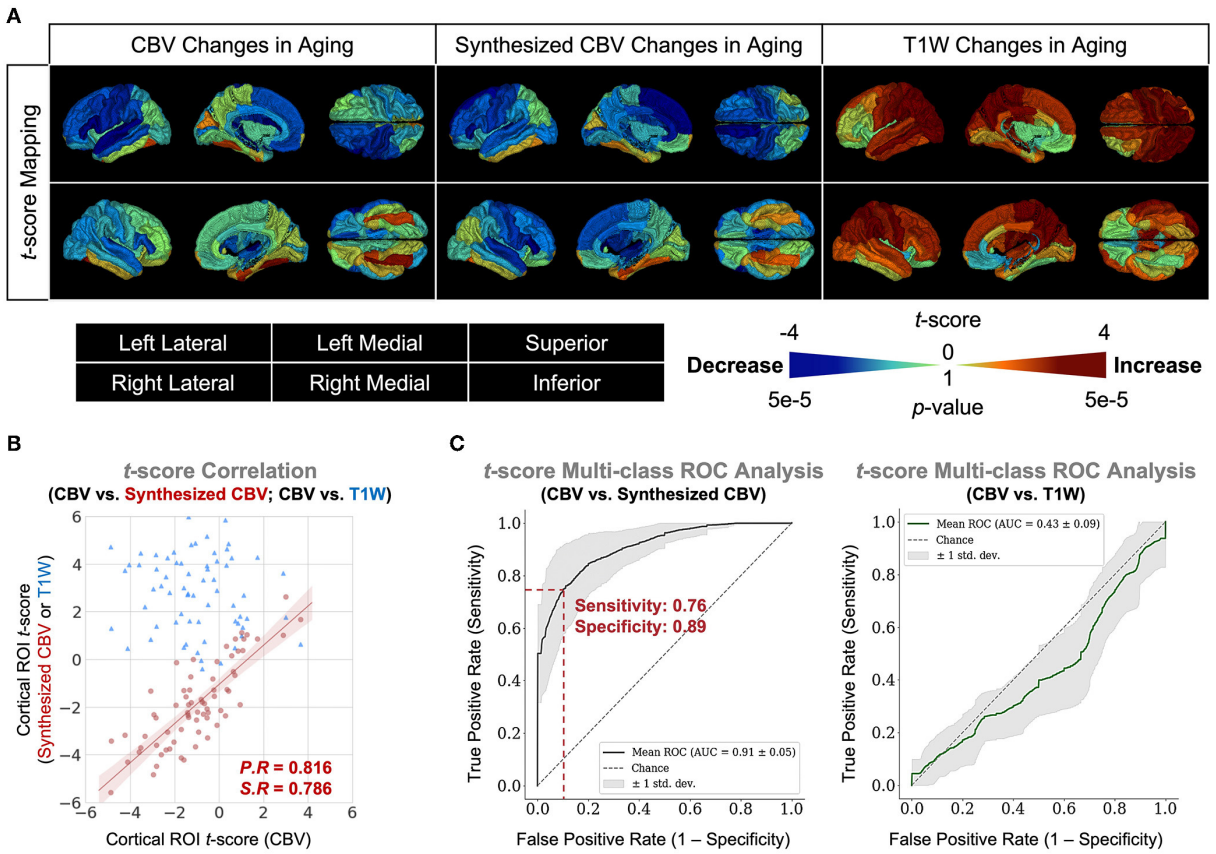


FIGURE 6
DeepContrast maps age-related changes over the entire cortex. **(A)** Three-dimensional volume rendering of the age-related *t*-score maps over the 72 FreeSurfer cortical region-of-interests (ROIs) reveals that the age-related changes in the synthesized CBV maps is similar to those in the ground truth CBV maps, despite that the non-contrast T1W scans which serves as the input to the DeepContrast model does not share either the same or the opposite trends. **(B)** A scatter plot of the age-related *t*-score over the 72 ROIs demonstrates that the age-related changes in synthesized CBV are consistent to those in CBV (P.R = 0.816, S.R = 0.786) while the T1W counterparts do not (P.R = −0.131, S.R = −0.122). **(C)** An analysis of the concordance to CBV *t*-scores by treating it as a 1,000-class classification problem reveals that age-related changes in synthesized CBV have significant predictive power on those in CBV (sensitivity = 0.76, specificity = 0.89, AUC = 0.91) while the T1W counterparts do not (sensitivity = 1.00, specificity = 0.04, AUC = 0.43).

right parahippocampal cortex (PHC). The region was defined as the intersection between the EC-PHC region and a sphere centered at the middle of the EC-PHC intersection and spanning a diameter of the extent of the EC-PHC boundary (11 mm). A box plot overlaid with individual data points was drawn (Figure 4G) to indicate the group-wise difference between the normal controls and the patients with AD.

Utility study: Synthesized CBV improves AD classification

For the AD classification tasks with one single input modality, the architecture "VGG-19 with batch normalization" (VGG-19BN) (Marcel et al., 2016) was used (Supplementary Figures S5a,b). When both T1W and synthesized CBV were used as input, each as one three-dimension(3D) volume, we used separate VGG encoders for each volume and later combined the extracted feature vectors before feeding them into fully-connected layers. The two encoders may have different weights (Supplementary Figure S5c). For any of these architectures, the input is the relevant 3D scans while the output is a continuous-valued number representing the predicted AD-likelihood.

To evaluate the descriptiveness of the predicted AD-likelihoods, receiver-operating characteristics (ROC) studies were conducted to analyze the concordance between the model-generated classification and the ground truth AD/CN labels. The ROC curves, one for each well-trained classifier, represent the classification performance at each potential numerical threshold to binarize the predicted AD-likelihood score. The sensitivity and specificity (the sum of whom peaks at the operating point), as well as the total area under the ROC curve, demonstrate the effectiveness of the classification method. The significance of the

difference among these ROC curves is calculated using DeLong's test (DeLong et al., 1988).

Furthermore, we investigated the brain regions that had the most contributions to the AD classification task by visualizing the class activation maps (CAM) (Bolei et al., 2015). We used all 131 T1W and 131 synthesized CBV scans from patients with AD to generate an averaged CAM for each input type. We were interested in whether or not the brain regions the classifier found most relevant to the AD class were in fact physiologically meaningful.

Results

DeepContrast in the mouse brain

We first designed, optimized, and trained the model on wildtype (WT) mice brain scans (37 for training and 6 for validation; refer to Methods section), in which we had previously generated quantitative T2-weighted GBCA-uptake brain maps. Similar to the previous studies (Kleesiek et al., 2019; Liu et al., 2019; Li et al., 2021), we compared the similarities between the GBCA-predicted maps and the GBCA-uptake ground truth maps by performing voxel-wise analyses across the whole brain on a test set with 6 scans (Figures 3A,B) using metrics that measure signal quality (peak signal-to-noise ratio) and structural similarity (structural similarity index). We further incorporated two other metrics to represent linear (Pearson correlation coefficient) and monotonic (Spearman correlation coefficient) relationships across corresponding voxels. Between the maps, the peak signal-to-noise ratio was 24.59 ± 0.60, the Pearson correlation coefficient was 0.695 ± 0.008 ($p < 0.0001$), the Spearman correlation coefficient was 0.606 ± 0.008 ($p < 0.0001$), and the structural similarity index was 0.831 ± 0.008 (Figure 3B and Table 1). This analysis shows that the DeepContrast-generated GBCA-predicted maps from WT mice showed high similarity to the GBCA-uptake ground truth maps generated from WT mice.

DeepContrast in the human brain

We adapted the DeepContrast model to human brain MRI datasets by modifying the model network architecture, hyper-parameters, and training strategies. First, same as in our mouse study, we compared the similarities between the GBCA-predicted images or maps, and the GBCA-uptake ground truth maps by performing voxel-wise analyses across the whole brain on a test set with 179 scans (Figures 3C,D). Between the maps, the peak signal-to-noise ratio was 29.64 ± 0.07, the Pearson correlation coefficient was 0.822 ± 0.002 ($p < 0.0001$), the Spearman correlation coefficient was 0.625 ± 0.003 ($p < 0.0001$), and the structural similarity index was 0.879 ± 0.002 (Figure 3D and Table 1). Thus, in healthy human brains, we also see a high similarity between the GBCA-predicted maps and that of the GBCA-uptake ground truth maps.

In addition to the whole-brain analysis for similarity measures, we decided to extend our comparisons to two additional analyses. In the tissue of interest (TOI) analysis, we compared the similarities between the maps in white matter, gray matter, and cerebrospinal fluid (CSF). Similar to the global results, the performances by tissue types demonstrated the same trend: the GBCA-predicted maps were quantitatively similar to the GBCA-uptake ground truth maps. The results are illustrated in Figure 3D and reported in Table 1. In the region-of-interest (ROI) analysis, we compared the similarities between the maps in 126 distinct ROIs in the whole brain segmented by FreeSurfer (Fischl, 2012). Among the 126 ROIs, 121 had a significant Pearson correlation coefficients ($p < 0.001$) and 123 had significant Spearman correlation coefficients ($p < 0.001$) (Supplementary Figure S3).

Finally, we were also interested in evaluating reproducibility in a test-retest paradigm. We conducted a series of test-retest reliability analyses on the GBCA-predicted maps vs. the GBCA-uptake ground truth maps across the whole brain on a test set with 11 repeated scan pairs (Figure 3E). For the GBCA-predicted maps, the peak signal-to-noise ratio was 30.11 ± 0.44, the Pearson correlation coefficient was 0.919 ± 0.005 ($p < 0.0001$), the Spearman correlation coefficient was 0.722 ± 0.010 ($p < 0.0001$), and the structural similarity index was 0.948 ± 0.002. As a comparison, for the GBCA-uptake ground truth maps, the peak signal-to-noise ratio was 27.70 ± 0.24, the Pearson correlation coefficient was 0.815 ± 0.007 ($p < 0.0001$), the Spearman correlation coefficient was 0.415 ± 0.017 ($p < 0.0001$), and the structural similarity index was 0.832 ± 0.007. Among all the analyses, the test-retest reliabilities of the GBCA-predicted maps were consistently higher than the test-retest reliabilities of the GBCA-uptake ground truth maps ($p < 0.0001$) (Figure 3E). We also performed the TOI analysis, and the results are illustrated in Figure 3E and reported in Table 1. Among all metrics in all tissue types, the test-retest reliabilities of the GBCA-predicted maps were consistently higher than the test-retest reliabilities of the GBCA-uptake ground truth maps ($p < 0.0001$) (Figure 3E).

DeepContrast visualizes functional lesions in aging and Alzheimer's disease brains

We generated GBCA-predicted maps from non-contrast T1-weighted MRI scans with DeepContrast, and subsequently quantified synthesized CBV maps with a sub-millimeter in-plane resolution of 0.68×0.68 mm in the coronal planes and slice thickness of 3 mm (refer to Methods). Then, we conducted

TABLE 1 Quantitative evaluations of DeepContrast models.

Model	Evaluation	Region	Data	PSNR	P.R	S.R	SSIM
Healthy Mouse	Similarity	Whole Brain	Gd-Uptake vs. T2W	22.16 ± 0.59	−0.072 ± 0.032	0.074 ± 0.013	0.707 ± 0.007
			Gd-Uptake vs. Gd-Predicted	24.59 ± 0.60	0.695 ± 0.008	0.606 ± 0.008	0.831 ± 0.008
Healthy Human	Similarity	Whole Brain	Gd-Uptake vs. T1W	15.40 ± 0.09	−0.194 ± 0.003	−0.323 ± 0.005	0.446 ± 0.002
			Gd-Uptake vs. Gd-Predicted	29.64 ± 0.07	0.822 ± 0.002	0.625 ± 0.003	0.879 ± 0.002
		White Matter	Gd-Uptake vs. T1W	15.40 ± 0.09	−0.194 ± 0.003	−0.323 ± 0.005	0.446 ± 0.002
			Gd-Uptake vs. Gd-Predicted	35.15 ± 0.09	0.633 ± 0.006	0.437 ± 0.002	0.969 ± 0.001
		Gray Matter	Gd-Uptake vs. T1W	14.72 ± 0.06	−0.029 ± 0.002	−0.039 ± 0.003	0.462 ± 0.001
			Gd-Uptake vs. Gd-Predicted	30.18 ± 0.07	0.807 ± 0.004	0.510 ± 0.003	0.946 ± 0.001
		CSF	Gd-Uptake vs. T1W	20.65 ± 0.07	0.124 ± 0.003	0.120 ± 0.004	0.802 ± 0.003
			Gd-Uptake vs. Gd-Predicted	25.47 ± 0.08	0.728 ± 0.004	0.604 ± 0.003	0.936 ± 0.001
	Test-Retest Reliability	Whole Brain	Gd-Uptake of two repeated acquisitions	27.70 ± 0.24	0.815 ± 0.007	0.415 ± 0.017	0.832 ± 0.007
			Gd-Predicted of two repeated acquisitions	30.11 ± 0.44	0.919 ± 0.005	0.722 ± 0.010	0.948 ± 0.002
		White Matter	Gd-Uptake of two repeated acquisitions	31.18 ± 0.18	0.713 ± 0.015	0.129 ± 0.008	0.933 ± 0.004
			Gd-Predicted of two repeated acquisitions	35.14 ± 0.40	0.899 ± 0.006	0.368 ± 0.014	0.986 ± 0.001
		Gray Matter	Gd-Uptake of two repeated acquisitions	27.89 ± 0.32	0.844 ± 0.008	0.327 ± 0.020	0.907 ± 0.005
			Gd-Predicted of two repeated acquisitions	30.42 ± 0.44	0.937 ± 0.004	0.596 ± 0.013	0.976 ± 0.001
		CSF	Gd-Uptake of two repeated acquisitions	24.04 ± 0.27	0.762 ± 0.014	0.585 ± 0.019	0.907 ± 0.004
			Gd-Predicted of two repeated acquisitions	26.32 ± 0.48	0.882 ± 0.009	0.800 ± 0.008	0.967 ± 0.001

Evaluations varied depending on the aspects being assessed for each model. All metrics were reported in the form of mean ± standard error of the mean (SEM). PSNR, peak signal-to-noise ratio; P.R, Pearson correlation coefficient; S.R, Spearman correlation coefficient; SSIM, structural similarity index.

voxel-based analyses (VBA) and ROI-based analyses on the synthesized CBV maps to identify sites of dysfunctions in normal aging and Alzheimer's disease (AD).

Normal aging

The first study we conducted aimed to validate whether DeepContrast can capture the subtle aging effects on basal metabolism. First, we focused on the hippocampal circuit (Figure 4A). As shown in Figures 4B,C, the age-related decline in our DeepContrast-generated synthesized CBV maps localized primarily to the dentate gyrus (DG). This result replicates prior studies (Small et al., 2002, 2004; Chawla and Barnes, 2007; Moreno et al., 2007; Brickman et al., 2014; Feng et al., 2020a), where the age-related decline in brain metabolism in the hippocampal formation has been shown to occur primarily in the DG. In the complementary ROI analysis of the DG, the synthesized CBV values showed a linear decline with age (β_{age} = −6.36e-4, t_{age} = −4.64, p_{age} = 6.85e-6) (Figure 4D).

We also analyzed two other brain regions, namely the inferior frontal gyrus (IFG), found to be more vulnerable to aging (Shen et al., 2012; Hoffman and Morcom, 2018; Feng et al., 2020a,b), and the entorhinal cortex (EC), found to be less vulnerable to aging (Gómez-Isla et al., 1996; Small et al., 2004, 2011; Feng et al., 2020a). The synthesized CBV maps demonstrated the same age-related trends as the ground truth CBV over these regions (Figure 5).

Finally, we extended the analysis to the entire cortex and found that the synthesized CBV maps reflected similar age-related changes as the ground truth CBV overall cortical ROIs (Figure 6). The multi-class Receiver Operative Characteristics (ROC) curve, which represented the level of concordance between synthesized CBV and ground truth CBV, reached a sensitivity of 0.76 and a specificity of 0.89 at the operating point, and the area-under-the-curve (AUC) was 0.91 (Figure 6).

Alzheimer's disease

The second study we conducted aimed to validate whether DeepContrast could capture the regional vulnerability in patients with Alzheimer's disease dementia, where we utilized publicly available data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) (Mueller et al., 2005). Similar to the above approach, we performed a VBA analysis over the hippocampal circuit (Figure 4A). Replicating previous findings (Braak and Braak, 1996; Brickman et al., 2011, 2014; Small et al., 2011; Pavlopoulos et al., 2013; Schobel et al., 2013; Khan et al., 2014; Sperling et al., 2014; Coughlan et al., 2018; Provenzano et al., 2020; Simoes et al., 2021), the Alzheimer's disease-related decline in the synthesized CBV maps primarily localized to a region termed the transentorhinal cortex (TEC) (Figures 4E,F). In the complementary ROI analysis of the right TEC, the synthesized CBV values were significantly

lower ($p = 0.031$) in patients with AD compared to the healthy controls (Figure 4G).

Next, we trained VGG-like models (Simon et al., 2016) to perform participant-level AD classification on a class-balanced and age-matched dataset with more than 2,500 scans. We tested the models on the same stand-alone set of 131 AD scans and 129 CN scans. Compared to using the T1W MRI data alone, when we included the synthesized CBV maps generated using DeepContrast as the input to the model, the classification accuracy increased significantly (Figure 7A). However, the specific approach to combining the two modalities affected the performances (Table 2). The best fit that we found was to assign a designated encoder for each modality without weight sharing between them.

Finally, we used class activation maps to identify specific brain regions that influenced the best performing classifier when it determined that a person has AD (Figure 7B). The most highly contributing structural information comes from the temporal lobe, while the most highly informative artificial functional information was observed to come from the parieto-occipital lobe.

Discussion

Gadolinium-based contrast agents's utility for MRI can be organized around two primary pathophysiologies. The first is a breakdown of the blood-brain barrier that often accompanies many structural lesions, and, in that case, GBCA extravasates into the parenchyma and enhances lesion detection (Garcia et al., 2020). The second is alterations in neuronal metabolism, typical of most functional disorders, in which case intravascular GBCA is used to quantify regional CBV, a cerebrovascular variable tightly coupled to energy metabolism (Belliveau, 1991; Gonz et al., 1995; Østergaard et al., 1998; Sugahara et al., 1998; Aronen et al., 2000). As proof-of-principle, we optimized models for our investigation of the second scenario. We have demonstrated that DeepContrast can extract GBCA-like contrast information from non-contrast T1W structural MRI scans to quantify regional CBV. As GBCA's utility can be largely divided into two pathophysiologies, we anticipate that future large-scale studies across a range of diseases might lead to two generalizable models–one for structural disorders that are more likely to impact the integrity of the blood-brain barrier, and another for functional lesions that alters brain metabolism, although GBCA contrast is much subtler for functional compared to structural lesions.

DeepContrast's utility can be organized according to its broad applications. The first is for research. There is an increasing number of brain MRI databases, such as ADNI, whose primary purpose is brain imaging and disease research. Standard T1-weighted MRI scans are among the most common protocols across all of these datasets, typically acquired for mapping regional structural differences, such as regional volume or cortical thickness. DeepContrast can be retroactively applied to these data and can be used to generate synthetic functional maps, significantly expanding pathophysiological insight that can be derived across a range of disorders. For example, by using the DeepContrast model, we have demonstrated that a large-scale synthetic functional dataset could be generated and further used to provide superior AD classification. For single input modalities, the AD classifier trained on the synthesized CBV functional images provided improved AD identification when compared with the AD classifier trained on T1W scans. The improvement was further amplified when both modalities were provided, which allowed the model to take advantage of both structural and functional information. Among all the candidate models, the model with two separate encoders of different weights outperforms the others. Training each encoder on structural and synthesize functional CBV MRI scans independently allowed the most efficient feature extraction and yielded the best classification performances. The class activation map for the best-performing model revealed an interesting pattern of collaboration between the two encoders, each corresponding to a single input modality. The medial temporal lobe provides the most crucial structural information as reflected by the structural-encoder. This result is consistent with previous studies indicating that medial temporal atrophy is an indicative sign of AD and qualitative assessments of the region could be used to predict patients at risk of AD (Bradley et al., 1984; Korf et al., 2004). On the other hand, activation of the parietal and occipital lobes was representative of regions experiencing the most functional changes in the AD brain in accordance with the functional-encoder, which is consistent with the findings such as decreased resting state neural activity (Yong et al., 2007; Li et al., 2016) and glucose utilization (Reiman et al., 1996) in the parieto-occipital cortex.

DeepContrast's second application is for patient care. For patient populations with functional lesions, those with neuropsychiatric and neurodegenerative disorders, a T1-weighted scan may be ordered as part of standard clinical practice, to exclude structural abnormalities. For these patients, deriving CBV maps*via* DeepContrast potentially obviates the need for ordering other more invasive, burdensome, and expensive neuroimaging studies for mapping metabolic dysfunction.

This study has several limitations. First, our study focused on the identification of functional abnormalities with the synthesized CBV derived from non-contrast T1W structural MRI without the need for GBCAs. Detecting structural brain lesions such as those found in cancer, infections, or bleeding are still the major utility of GBCAs. Our DeepContrast framework is sufficiently general that it can be easily extended to the detection of structural brain lesions with T1W MRIs, but future study should be done to evaluate its performance in these conditions. Second, we acknowledge the retrospective

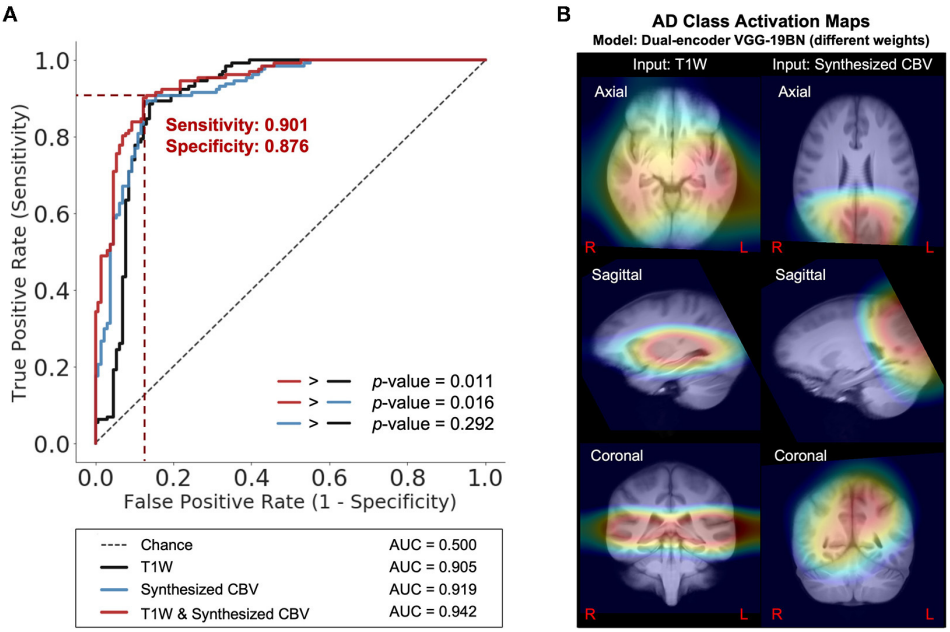


FIGURE 7
DeepContrast-generated synthesized CBV improves AD classification and provides spatially-meaningful information. **(A)** Receiver Operating Characteristics curves for the classification models using T1W structural data only (black), using synthesized CBV functional data only (blue), or using both in combination (red). *p*-values were calculated using DeLong's test. **(B)** The class-average Class Activation Maps of the best-performing model, calculated from 131 AD scans in the standalone test set.

TABLE 2 Performances of the five variants of the AD-classification network.

Input	Model	Sensitivity	Specificity	ROC AUC	Accuracy @ Operating	Max accuracy
T1W	Regular	0.885	0.860	0.905	0.869	0.873
SynCBV	Regular	0.885	**0.876**	0.919	0.877	0.881
T1W+SynCBV	Dual-channel	0.885	0.853	0.936	0.865	0.869
T1W+SynCBV	Dual-encoder w/ identical weights	0.802	0.806	0.875	0.800	0.804
T1W+SynCBV	Dual-encoder w/ different weights	**0.901**	**0.876**	**0.942**	**0.885**	**0.888**

Sensitivity and specificity are calculated at the operating point. Accuracy at the operating point and the maximum accuracy achievable by changing the binarization threshold are respectively calculated for each candidate. ROC AUC, area under the receiver-operating characteristics curve; SynCBV, synthesized CBV. Best result(s) in each metric are highlighted in bold.

nature of the study and the absence of inclusion of MRI data from multiple sites and acquisition protocols. The distinct quantitative GBCA dataset used to train our DeepContrast model was collected over a timeframe of around 20 years at Columbia University using the equivalent protocol on multiple scanners. The pre- and post-GBCA images used to derive the CBV maps shared identical imaging settings, which help reduce the inter and intrasubject variability between the non-contrast T1W pre-GBCA image and the quantitative GBCA enhancement and benefit the training of our model. Future study could address how the use of heterogeneous data from various cohorts, sites, scanners, and acquisition protocols might improve the model performance to produce more stable and generalizable results. This study also shares the limitations of other studies of GBCA contrast synthesizing with deep learning. Deep convolutional neural networks have performed remarkably well on these tasks; however, these networks are heavily reliant on big data to avoid overfitting. Unfortunately, medical image analysis applications normally do not have access to big data. Data augmentation, a data-space solution to the problem, of limited data can be applied in future work to enhance the sample size, enrich the data variance and improve the data quality of the training dataset such that better models can be built from them. Finally, in

light of the overall promising performance on the task, it is important to consider our work as a pioneer proof-of-concept study, and future work should be done to further improve its performance to reach the level for clinical purposes. For instance, the proposed CNN model is trained from scratch, but applying and fine-tuning a pre-trained model on our data through transfer learning could further improve model performance. Future study could also improve our model through multi-task learning. While we did achieve state-of-the-art performance by being laser-focused on our single task, synthesizing CBV from non-contrast T1W structural MRI, we ignored information that might help us achieve even better metrics. Specifically, this information could come from training MRI signals for related tasks, such as image reconstruction, brain tissue segmentation, or predicting demographic information. By sharing representations between related tasks in a multi-task learning framework, it could enable our model to generalize better on the original task.

In conclusion, by using quantitative GBCA datasets from both mice and humans, we demonstrated that a deep learning model can, in principle, generate GBCA-equivalent information from a single structural MRI scan for the estimation of regional CBV, and we successfully applied our DeepContrast model to both an in-house aging dataset and a publicly available ADNI dataset from Alzheimer's patients and age-matched controls.

Data availability statement

The trained Healthy Human Model, alongside the test-retest reliability dataset (n = 11, each with two test-retest acquisitions) with both non-contrast scans and ground truth GBCA-uptake maps, is available on GitHub (https://github.com/SAIL-GuoLab/DeepContrast_Demo). The scripts that predict GBCA-uptake maps from non-contrast scans, as well as the script performing quantitative evaluations, are included. All code and data (except for those from public datasets) are proprietary and managed by the Columbia Technology Ventures Office of Intellectual Property. The custom training code and large-scale datasets are not publicly available. Further inquiries can be directed to the corresponding author.

Ethics statement

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee (IACUC), NIH.

Author contributions

JG conceived, designed, and supervised all studies constituting this article and verified all statistical results. CL and NZ participated in the design and optimization of all DeepContrast models, and particularly optimized and trained the healthy human model including data preprocessing, performed statistical analyses of the Healthy Human Model, conducted the Human Aging study, and conducted the Human AD study. HS and XL led the optimization and training of the Healthy Mouse Model including data preprocessing. SS kindly provided all human MRI data in the Columbia University cohort, which was collected, organized, and maintained by FP and XF kindly provided the organized and preprocessed ADNI data, and the brain parcellations of the 177 healthy participants used in the Human Aging study. JG and SG-S participated in the acquisition of the mouse MRI data. HS performed statistical analyses of the healthy mouse model. CL, NZ, SS, FP, TN, and JG wrote the manuscript. NZ, CL, and JZ created and updated all display items (figures and tables) and supplementary information (figures and tables). All authors reviewed, commented, and edited the manuscript. Data used in the Human AD study was obtained from the ADNI database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this manuscript.

Funding

This study was funded by the Seed Grant Program and Technical Development Grant Program at the Columbia MR Research Center. This study was also funded by grants from Alzheimer's Disease Research Center (P30AG066462), Matheson Foundation (UR010590), and the cancer center support grant (P30CA013696 NIH/NIC). This study was performed at the Zuckerman Mind Brain Behavior Institute MRI Platform, a shared resource. Data collection and sharing for this project was partially funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada.

Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Conflict of interest

Author FP is a consultant for and equity holder of Imij Technologies. Author SS serves on the scientific advisory board of Meira GTX, recently came off the scientific advisory board of Denali Theraputics, and is an equity holder in Imij Technologies. Authors XF, FP, SS, and JG have either granted patents or applications in neuroimaging for which no royalties are received.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.923673/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Details of the Healthy Mouse Model (Network 1 in Figure 2). This model is implemented with a 2D five-layered Residual Attention U-Net architecture. The encoding path (blue blocks in the left half of the architecture) condenses the image dimensions and enriches the feature dimension, shrinking the image size from 208×208 pixels to 13×13 pixels while extracting 1,024 channels of features. The decoding path (red blocks in the right half of the architecture) expands these high-level features and returns back a single slice of the predicted image with 208×208 pixels.

SUPPLEMENTARY FIGURE S2
Details of the Healthy Human Model (Network 2 in Figure 2). This model is implemented with a 2D six-layered Residual Attention U-Net architecture. The encoding path (blue blocks in the left half of the architecture) condenses the image dimensions and enriches the feature dimension, shrinking the image size from 352×352 pixels to 11×11 pixels while extracting 2,048 channels of features. The decoding path (red blocks in the right half of the architecture) expands these high-level features and returns back a single slice of the predicted image with 352×352 pixels.

SUPPLEMENTARY FIGURE S3
Correlation between the GBCA-uptake map and GBCA-predicted map over the entire brain. The Pearson (top) and Spearman (bottom) correlation coefficients are computed for each of the 126 anatomical ROIs defined by FreeSurfer between the array of ROI-mean Gd-Uptake values and the array of ROI-mean Gd-Predicted values extracted from the 177 participants with successful FreeSurfer parcellation. The analysis demonstrates significant correlation between the DeepContrast-predicted contrast and the corresponding ground truth across a wide range of brain regions (121 ROIs with $p < 0.001$ for Pearson correlation coefficient and 123 ROIs with $p < 0.001$ for Spearman correlation coefficient).

SUPPLEMENTARY FIGURE S4
Details of MRI processing pipeline and partition of the Alzheimer's Disease (AD) classification dataset. **(a)** Data processing pipeline to generate the input of different AD classification deep learning models. The preprocessing of structural T1W MR data is necessary to remove unwanted artifacts and transform the data into a standard version before feeding them into the models. For each MRI structural, we process the T1W 3D volume through a standardized pipeline: (1) whole brain affine registration to MNI152 template space and other processing detailed in a prior study (Feng et al., 2018); (2) whole brain rigid registration to MNI152 template space; (3) histogram matching to the DeepContrast CU T1W MRI; (4) whole brain diffeomorphic registration to standard CU template; (5) generating the synthesized CBV maps; (6) up-sampling the synthesized CBV maps to 1-mm isotropic resolution. **(b)** Left: Age distributions of the participants in the entire dataset (top) and the subset after propensity-score matching of age (bottom). Right: Age distributions of the participants assigned to the train, validation, and test dataset. VS, voxel size; AD, Alzheimer's diseased; CN, cognitive normal.

SUPPLEMENTARY FIGURE S5
Details of the three candidates of AD-classification networks (Networks used in Figure 7). **(a)** Proposed Model 1. Modified 3D VGG-19 network with batch normalization uses T1W MRI scans as the model input. This model aims to perform the AD classification based on structure T1W MRI scans. **(b)** Proposed Model 2. Modified 3D VGG-19 network with batch normalization uses synthesized CBV maps as the model input. This model aims to realize the AD classification based on functional MRI maps. **(c)** Proposed model 3. Modified VGG-19 network with double encoding paths uses T1W MRI scans and the synthesized CBV maps as the model input. Separate encoders in the model have the same structure but different weights. This could help improve the structure and functional information fusion and classification performance. The red arrow represents the generation of synthesized CBV from T1W MRI scans using DeepContrast.

References

Aronen, H. J., Pardo, F. S., Kennedy, D. N., Belliveau, J. W., Packard, S. D., Hsu, D. W., et al. (2000). High microvascular blood volume is associated with high glucose uptake and tumor angiogenesis in human gliomas. *Clin. Cancer Res.* 6, 2189–2200. Available online at: https://aacrjournals.org/clincancerres/article/6/6/2189/288412/High-Microvascular-Blood-Volume-Is-Associated-with

Ashburner, J., Barnes, G., Chen, C. C., Daunizeau, J., Flandin, G., Friston, K., et al. (2014). *Spm12 Manual*. London: Wellcome Trust Centre for Neuroimaging.

Avants, B. B., and Tustison, N., and Song, G. (2009). Advanced normalization tools (ants). *Insight J.* 2, 1–35. doi: 10.54294/uvnhin

Barron, J. T. (2019). "A General and adaptive robust loss function," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Long Beach, CA: IEEE), 4331–4339.

Belliveau, J., et al. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging. *Science* 254, 716–719. doi: 10.1126/science.1948051

Bolei, A., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A. (2015). Learning deep features for discriminative localization. *arXiv:1512, 04150*. doi: 10.48550/arXiv.1512.04150

Borges, A., López-Larrubia, P., Marques, J., and Cerdan, S. (2012). Mr imaging features of high-grade gliomas in murine models: how they compare with human disease, reflect tumor biology, and play a role in preclinical trials. *Am. J. Neuroradiol*. 33, 24–36. doi: 10.3174/ajnr.A2959

Braak, H., and Braak, E. (1996). Evolution of the neuropathology of alzheimer's disease. *Acta Neurol Scand*. 94, 3–12. doi: 10.1111/j.1600-0404.1996.tb05866.x

Bradley, C., Van Hoesen, G. W., Damasio, A. R., and Barnes, C. L. (1984). Alzheimer's disease: cell-specific pathology isolates the hippocampal formation. *Science* 225, 1168–1170. doi: 10.1126/science.6474172

Brickman, A. M., Khan, U. A., Provenzano, F. A., Yeung, L. K., Suzuki, W., Schroeter, H., et al. (2014). Enhancing dentate gyrus function with dietary flavanols improves cognition in older adults. *Nat. Neurosci*. 17, 1798–1803. doi: 10.1038/nn.3850

Brickman, A. M., and Stern, Y., and Small, S. A. (2011). Hippocampal subregions differentially associate with standardized memory tests. *Hippocampus* 21, 923–928. doi: 10.1002/hipo.20840

Chawla, M. K., and Barnes, C. A. (2007). Hippocampal granule cells in normal aging: insights from electrophysiological and functional imaging experiments. *Prog. Brain Res*. 163, 661–821. doi: 10.1016/S0079-6123(07)63036-2

Chou, N., Wu, J., Bingren, J. B., and Qiu, A., and Chuang, K.-H. (2011). Robust automatic rodent brain extraction using 3-d pulse-coupled neural networks (pcnn). *IEEE Trans. Image Process*. 20, 2554–2564. doi: 10.1109/TIP.2011.2126587

Coughlan, G., Laczó, J., Hort, J., Minihane, A.-M., and Hornberger, M. (2018). Spatial navigation deficits–overlooked cognitive marker for preclinical alzheimer disease? *Nat. Rev. Neurol*. 14, 496–506. doi: 10.1038/s41582-018-0031-x

Cox, I. J., and Roy, S., and Hingorani, S. L. (1995). "Dynamic histogram warping of image pairs for constant image brightness," in *Proceedings., International Conference on Image Processing, Vol. 2* (Washington, DC: IEEE), 366–369.

Cox, R. W. (1996). Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res*. 29, 162–173. doi: 10.1006/cbmr.1996.0014

Cox, R. W., Chen, G., Glen, D. R., and Reynolds, R. C., and Taylor, P. A. (2017). Fmri clustering in afni: false-positive rates redux. *Brain Connect*. 7, 152–171. doi: 10.1089/brain.2016.0475

DeLong, D., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 33, 837–845. doi: 10.2307/2531595

Dillman, J. R., and Davenport, M. S. (2020). Gadolinium retention–5 years later. *Pediatr. Radiol*. 50, 166–167. doi: 10.1007/s00247-019-04540-z

Dugas-Phocion, G., Ballester, M. A. G., Malandain, G., Lebrun, C., and Ayache, N. (2004). "Improved EM-based tissue segmentation and partial volume effect quantification in multi-sequence brain MRI," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* Berlin.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., J. Fillion-Robin, C., Pujol, S., et al. (2012). 3D slicer as an image computing platform for the quantitative imaging network. *Magn. Reson. Imaging* 30, 1323–1341. doi: 10.1016/j.mri.2012.05.001

Feng, X., Guo, J., Sigmon, H. C., Sloan, R. P., Brickman, A. M., Provenzano, F. A., et al. (2020a). Brain regions vulnerable and resistant to aging without alzheimer's disease. *PLoS ONE* 15, e0234255. doi: 10.1371/journal.pone.0234255

Feng, X., Lipton, Z. C., Yang, J., Small, S. A., Provenzano, F. A., Alzheimer's Disease Neuroimaging Initiative, et al. (2020b). Estimating brain age based on a uniform healthy population with deep learning and structural mri. *Neurobiol. Aging* 91, 15–25. doi: 10.1016/j.neurobiolaging.2020.02.009

Feng, X., Yang, J., Lipton, Z. C., Small, S. A., and Provenzano, F. A. (2018). Deep learning on MRI affirms the prominence of the hippocampal formation in Alzheimer's disease classification. *bioRxiv* 456277. doi: 10.1101/456277

Fischl, B. (2012). Freesurfer. *Neuroimage* 62, 774–781. doi: 10.1016/j.neuroimage.2012.01.021

Flacke, S., Urbach, H., Keller, E., Träber, F., Hartmann, A., Folkers, P. J., et al. (2000). Middle cerebral artery (MCA) susceptibility sign at susceptibility-based perfusion MR imaging: clinical importance and comparison with hyperdense MCA sign at CT. *Radiology* 215, 476–482. doi: 10.1148/radiology.215.2.r00ma09476

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., and Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fmri): use of a cluster-size threshold. *Magn. Reson. Med*. 33, 636–647. doi: 10.1002/mrm.1910330508

Garcia, G. C., Bockel, S., Majer, M., and Ammari, S., and Smits, M. (2020). "Imaging of brain metastases: diagnosis and monitoring," in *Central Nervous System Metastases* (Springer), 145–158.

Gómez-Isla, T., Price, J. L., McKeel Jr, D. W., Morris, J. C., Growdon, J. H., and Hyman, B. T. (1996). Profound loss of layer ii entorhinal cortex neurons occurs in very mild Alzheimer's disease. *J. Neurosci*. 16, 4491–4500. doi: 10.1523/JNEUROSCI.16-14-04491.1996

Gong, E., Pauly, J. M., and Wintermark, M., and Zaharchuk, G. (2018). Deep learning enables reduced gadolinium dose for contrast-enhanced brain mri. *J. Magn. Reson. Imaging* 48, 330–340. doi: 10.1002/jmri.25970

Gonz, R., Fischman, A. J., Guimaraes, A. R., Carr, C. A., Stern, C. E., Halpern, E. F., et al. (1995). Functional mr in the evaluation of dementia: correlation of abnormal dynamic cerebral blood volume measurements with changes in cerebral metabolism on positron emission tomography with fludeoxyglucose f 18. *Am. J. Neuroradiol*. 16, 1763–1770.

Guo, B. J., and Yang, Z. L., and Zhang, L. J. (2018). Gadolinium deposition in brain: current scientific evidence and future perspectives. *Front. Mol. Neurosci*. 11, 335. doi: 10.3389/fnmol.2018.00335

Hasgall, P. A., Gennaro, D., i., Baumgartner, F., Neufeld, C., Lloyd, E., et al. (2022). *IT'IS Database for thermal and electromagnetic parameters of biological tissues*. Version 4, 1.

He, K., Zhang, X., and Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV: IEEE), 770–778.

Hoffman, P., and Morcom, A. M. (2018). Age-related changes in the neural networks supporting semantic cognition: a meta-analysis of 47 functional neuroimaging studies. *Neurosci. Biobehav. Rev*. 84, 134–150. doi: 10.1016/j.neubiorev.2017.11.010

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). FSL. *Neuroimage* 62, 782–790. doi: 10.1016/j.neuroimage.2011.09.015

Khan, U. A., Liu, L., Provenzano, F. A., Berman, D. E., Profaci, C. P., Sloan, R., et al. (2014). Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nat. Neurosci*. 17, 304–311. doi: 10.1038/nn.3606

Kleesiek, J., Morshuis, J. N., Isensee, F., Deike-Hofmann, K., Paech, D., Kickingereder, P., et al. (2019). Can virtual contrast enhancement in brain mri replace gadolinium?: a feasibility study. *Invest Radiol*. 54, 653–660. doi: 10.1097/RLI.0000000000000583

Korf, P., Wahlund, L. O., Visser, P. J., and Scheltens, P. (2004). Medial temporal lobe atrophy on MRI predicts dementia in patients with mild cognitive impairment. *Neurology* 63, 94–100. doi: 10.1212/01.WNL.0000133114.92694.93

Lewandowski, N. M., Bordelon, Y., Brickman, A. M., Angulo, S., Khan, U., Muraskin, J., et al. (2013). Regional vulnerability in huntington's disease: fmri-guided molecular analysis in patients and a mouse model of disease. *Neurobiol. Dis*. 52, 84–93. doi: 10.1016/j.nbd.2012.11.014

Li, H., Wang, X., Li, Y., Sun, Y., Sheng, C., Li, H., et al. (2016). Abnormal resting-state functional connectivity strength in mild cognitive impairment and its conversion to Alzheimer's disease. *Neural Plast*. 2016, 4680972. doi: 10.1155/2016/4680972

Li, W., Xiao, H., Li, T., Ren, G., Lam, S., Teng, X., et al. (2021). A multimodality-guided synergistic neural network (MMgSN-Net) was developed to leverage complementary information between contrast-free T1-weighted and T2-weighted MRI for vceT1w MRI synthesis.. *Int. J. Radiat. Oncol. Biol. Phys*. 112, 1033–1044.doi: 10.1016/j.ijrobp.2021.11.007

Liu, J., Gong, E., and Christen, T., and Zaharchuk, G. (2019). "Contrast-free mri contrast enhancement with deep attention generative adversarial network," in *ISMRM 27th Annual Meeting and Exhibition* Montréal, QC.

Lohrke, J., Frenzel, T., Endrikat, J., Alves, F. C., Grist, T. M., Law, M., et al. (2016). 25 years of contrast-enhanced mri: developments, current challenges and future perspectives. *Adv. Ther*. 33, 1–28. doi: 10.1007/s12325-015-0275-4

Marcel, J., Rodner, E., and Denzler, J. (2016). ImageNet pre-trained models with batch normalization. *arXiv:1612, 01452*. doi: 10.48550/arXiv.1612.01452

Moreno, H., Hua, F., and Brown, T., and Small, S. (2006). Longitudinal mapping of mouse cerebral blood volume with mri. *NMR Biomed*. 19, 535–543. doi: 10.1002/nbm.1022

Moreno, H., Wu, W. E., Lee, T., Brickman, A., Mayeux, R., Brown, T. R., et al. (2007). Imaging the aβ-related neurotoxicity of Alzheimer disease. *Arch. Neurol.* 64, 1467–1477. doi: 10.1001/archneur.64.10.1467

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C., Jagust, W., et al. (2005). The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clin.* 15, 869–877. doi: 10.1016/j.nic.2005.09.008

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., et al. (2018). Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804, 03999*. doi: 10.48550/arXiv.1804.03999

Ostergaard, L., Smith, D. F., Vestergaard-Poulsen, P., Hansen, S. B., Gee, A. D., Gjedde, A., et al. (1998). Absolute cerebral blood flow and blood volume measured by magnetic resonance imaging bolus tracking: comparison with positron emission tomography values. *J. Cereb. Blood Flow Metabol.* 18, 425–432. doi: 10.1097/00004647-199804000-00011

Pavlopoulos, E., Jones, S., Kosmidis, S., Close, M., Kim, C., Kovalerchik, O., et al. (2013). Molecular mechanism for age-related memory loss: the histone-binding protein rbap48. *Sci. Transl. Med.* 5, 200ra.115–200ra115. doi: 10.1126/scitranslmed.3006373

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830. Available online at: https://dl.acm.org/doi/10.5555/1953048.2078195 doi: 10.48550/arXiv.1201.0490

Preetha, C. J., Meredig, H., Brugnara, G., Mahmutoglu, M. A., Foltyn, M., Isensee, F., et al. (2021). Deep-learning-based synthesis of post-contrast T1-weighted MRI for tumour response assessment in neuro-oncology: a multicentre, retrospective cohort study. *Lancet Digital Health* 3, e784-e794. doi: 10.1016/S2589-7500(21)00205-3

Provenzano, F. A., Guo, J., Wall, M. M., Feng, X., Sigmon, H. C., Brucato, G., et al. (2020). Hippocampal pathology in clinical high-risk patients and the onset of schizophrenia. *Biol. Psychiatry* 87, 234–242. doi: 10.1016/j.biopsych.2019.09.022

Quattrocchi, C. C., and van der Molen, A. J. (2017). Gadolinium retention in the body and brain: is it time for an international joint research effort?. *Radiology* 282, 12–16. doi: 10.1148/radiol.2016161626

Ramalho, M., Ramalho, J., and Burke, L. M., and Semelka, R. C. (2017). Gadolinium retention and toxicity–an update. *Adv. Chronic Kidney Dis.* 24, 138–146. doi: 10.1053/j.ackd.2017.03.004

Reiman, E. M., Caselli, R. J., Yun, L. S., Chen, K., Bandy, D., Minoshima, S., et al. (1996). Preclinical evidence of Alzheimer's disease in persons homozygous for the ε4 allele for apolipoprotein E. *N. Engl. J. Med.* 334, 752–758. doi: 10.1056/NEJM199603213341202

Ronneberger, O., and Fischer, P., and Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Munich: Springer), 234–241.

Rovira, A., Orellana, P., Alvarez-Sabín, J., Arenillas, J. F., Aymerich, X., Grivé, E., et al. (2004). Hyperacute ischemic stroke: middle cerebral artery susceptibility sign at echo-planar gradient-echo MR imaging. *Radiology* 232, 466–473. doi: 10.1148/radiol.2322030273

Schobel, S. A., Chaudhury, N. H., Khan, U. A., Paniagua, B., Styner, M. A., Asllani, I., Inbar, B. P., et al. (2013). Imaging patients with psychosis and a mouse model establishes a spreading pattern of hippocampal dysfunction and implicates glutamate as a driver. *Neuron* 78, 81–93. doi: 10.1016/j.neuron.2013.02.011

Shen, Q., and Duong, T. Q. (2016). Magnetic resonance imaging of cerebral blood flow in animal stroke models. *Brain Circ.* 2, 20. doi: 10.4103/2394-8108.178544

Shen, X., Liu, H., Hu, Z., and Hu, H., and Shi, P. (2012). The relationship between cerebral glucose metabolism and age: report of a large brain pet data set. *PLoS ONE* 7, e51517. doi: 10.1371/journal.pone.0051517

Simoes, S., Guo, J., Buitrago, L., Qureshi, Y. H., Feng, X., Kothiya, M., et al. (2021). Alzheimer's vulnerable brain region relies on a distinct retromer core dedicated to endosomal recycling. *Cell Rep.* 37, 110182. doi: 10.1016/j.celrep.2021.110182

Simon, M., Rodner, E., and Denzler, J. (2016). Imagenet pre-trained models with batch normalization. *arXiv preprint arXiv:1612, 01452*.

Small, S. A., Chawla, M. K., Buonocore, M., and Rapp, P. R., and Barnes, C. A. (2004). Imaging correlates of brain function in monkeys and rats isolates a hippocampal subregion differentially vulnerable to aging. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7181–7186. doi: 10.1073/pnas.0400285101

Small, S. A., Schobel, S. A., Buxton, R. B., and Witter, M. P., and Barnes, C. A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nat. Rev. Neurosci.* 12, 585–601. doi: 10.1038/nrn3085

Small, S. A., and Tsai, W. Y., DeLaPaz, R., Mayeux, R., and Stern, Y. (2002). Imaging hippocampal function across the human life span: is memory decline normal or not? *Ann. Neurol.* 51, 290–295. doi: 10.1002/ana.10105

Small, S. A., Wu, E. X., Bartsch, D., Perera, G. M., Lacefield, C. O., DeLaPaz, R., et al. (2000). Imaging physiologic dysfunction of individual hippocampal subregions in humans and genetically modified mice. *Neuron* 28, 653–664. doi: 10.1016/S0896-6273(00)00144-6

Sperling, R., and Mormino, E., and Johnson, K. (2014). The evolution of preclinical Alzheimer's disease: implications for prevention trials. *Neuron* 84, 608–622. doi: 10.1016/j.neuron.2014.10.038

Sugahara, T., Korogi, Y., Kochi, M., Ikushima, I., Hirai, T., Okuda, T., et al. (1998). Correlation of mr imaging-determined cerebral blood volume maps with histologic and angiographic determination of vascularity of gliomas. *AJR Am. J. Roentgenol.* 171, 1479–1486. doi: 10.2214/ajr.171.6.9843274

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). "Attention is all you need," in *Advances in Neural Information Processing Systems* (Long Beach, CA), 5998–6008.

Wagenknecht, G., Kaiser, H.-J., Sabri, O., and Buell, U. (2000). "Dynamic programming algorithm for contrast correction in medical images," in *Nonlinear Image Processing XI, Vol. 3961* (San Jose, CA: International Society for Optics and Photonics), 216–226.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). "Image quality assessment: from error visibility to structural similarity," in *ProceedingsIEEE Transactions on Image Processing*, 13, 600–612.

Yong, T., Wang, L., Zang, Y., Tian, L., Zhang, X., Li, K., et al. (2007). Regional coherence changes in the early stages of Alzheimer's disease: a combined structural and resting-state functional MRI study. *Neuroimage* 35, 488–500. doi: 10.1016/j.neuroimage.2006.11.042

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 16 August 2022
DOI 10.3389/fnagi.2022.925227

OPEN ACCESS

EDITED BY
Wai Haung Yu,
University of Toronto, Canada

REVIEWED BY
Zeneng Wang,
Cleveland Clinic, United States
Claudia Saraiva,
University of Luxembourg,
Luxembourg

*CORRESPONDENCE
Mi-Ok Lee
molee@kribb.re.kr
Mi-Young Son
myson@kribb.re.kr

SPECIALTY SECTION
This article was submitted to Cellular and Molecular Mechanisms of Brain-aging, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 21 April 2022
ACCEPTED 18 July 2022
PUBLISHED 16 August 2022

CITATION
Lee Y, Kang JS, Ham O-J, Son M-Y and Lee M-O (2022) Gut metabolite trimethylamine *N*-oxide induces aging-associated phenotype of midbrain organoids for the induced pluripotent stem cell-based modeling of late-onset disease.
Front. Aging Neurosci. 14:925227.
doi: 10.3389/fnagi.2022.925227

Gut metabolite trimethylamine *N*-oxide induces aging-associated phenotype of midbrain organoids for the induced pluripotent stem cell-based modeling of late-onset disease

Youngsun Lee[1,2], Ji Su Kang[1,2], On-Ju Ham[1], Mi-Young Son[1,2]* and Mi-Ok Lee[1,2]*

[1]Stem Cell Convergence Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, South Korea, [2]Department of Bioscience, Korea University of Science and Technology (UST), Daejeon, South Korea

Brain organoids are valuable research models for human development and disease since they mimic the various cell compositions and structures of the human brain; however, they have challenges in presenting aging phenotypes for degenerative diseases. This study analyzed the association between aging and the gut metabolite trimethylamine *N*-oxide (TMAO), which is highly found in the midbrain of elderly and Parkinson's disease (PD) patients. TMAO treatment in midbrain organoid induced aging-associated molecular changes, including increased senescence marker expression (*P21, P16*), p53 accumulation, and epigenetic alterations. In addition, TMAO-treated midbrain organoids have shown parts of neurodegeneration phenotypes, including impaired brain-derived neurotrophic factor (BDNF) signaling, loss of dopaminergic neurons, astrocyte activation, and neuromelanin accumulation. Moreover, we found TMAO treatment-induced pathophysiological phosphorylation of α-synuclein protein at Ser-129 residues and Tau protein at Ser202/Thr205. These results suggest a role of TMAO in the aging and pathogenesis of the midbrain and provide insight into how intestinal dysfunction increases the risk of PD. Furthermore, this system can be utilized as a novel aging model for induced pluripotent stem cell (iPSC)-based modeling of late-onset diseases.

KEYWORDS

brain organoid, midbrain, gut metabolite, TMAO, aging, neurodegenerative disease

Introduction

Aging is the degenerative alterations in the body with age, accompanied by physiological, behavioral, and psychological changes, and contributes to the onset of diseases. In particular, aging is closely related to neurodegeneration such as Parkinson's disease (PD) in the central nervous system (CNS). Although various genetic and environmental factors that cause PD have been reported, aging is the greatest risk factor for PD. However, it remains unknown which specific mechanisms of aging induce the pre-parkinsonian state or what other mechanisms isolated from aging lead to PD (Pang et al., 2019). Understanding the detailed molecular mechanism of aging-related alteration and disease onset, especially in the human brain, will be a breakthrough in discovering treatments to cure.

Although numerous animal models are being studied to elucidate the relationship between aging and PD and identify the causes of diseases, these models face some challenges, including having different anatomical, and physiological characteristics from humans (Potashkin et al., 2010). As a method of studying human diseases by more closely mimicking the human microenvironment, induced pluripotent stem cell (iPSC) technology has recently been in the spotlight (Liu et al., 2018). In particular, brain organoids that can be produced through three-dimensional neuronal differentiation of human pluripotent stem cells (hPSCs) provide an advanced *in vitro* model for studying human brain development and disease by mimicking the cell types and tissue architecture of the human brain (Koo et al., 2019). However, since cellular reprogramming is known to erase aging-related features of patient cells, such as DNA damage, mitochondrial reactive oxygen species (ROS), nuclear envelope dysfunction, and epigenetic alteration (Simpson et al., 2021), it is limited to recapitulating the alterations of neurodegenerative diseases that show late-onset phenotypes with aging in induced pluripotent stem cell-based disease models. For instance, an increase in α-synuclein could be observed in an *in vitro* PD model using patient iPSC-derived midbrain organoids; however, Lewy body-like inclusions, a hallmark of PD, were not formed (Kim et al., 2019). This is why we need a model showing aging factors for PD study.

Given the importance of degenerative brain disease research, many studies are being conducted on methods to simulate them (Brunet, 2020; Slanzi et al., 2020). In iPSC-based studies, modeling for the implementation of an aging model including a progeria-induced model is being studied (Miller et al., 2013). However, the method through genetic modification is difficult to reproduce the phenomenon of natural aging (Azam et al., 2021). As a factor inducing such natural aging, the field of the gut microbiome is receiving attention recently (Ghosh et al., 2022). The gut microbiota plays an important role in various human health and diseases ranging from the immune system, metabolic disorders, and cancer, and shows potential as a biomarker (Zhu et al., 2020). Several studies have reported that gastrointestinal dysfunction is associated with PD risk (Fasano et al., 2015; Mukherjee et al., 2016; Schapira et al., 2017; Metta et al., 2022). Microbiota dysbiosis is related not only to atherosclerosis and stroke in the brain, but also to neurodegeneration such as Alzheimer's disease (AD), autism, multiple sclerosis, and PD (Janeiro et al., 2018; Gandy et al., 2019; Glowacki and Martens, 2020). Based on the latest report on the formation of a unique microbiome with aging (Wilmanski et al., 2021), the aging-related microbiome may be involved in the pathogenesis of late-onset diseases, including PD.

Trimethylamine *N*-oxide (TMAO) is a metabolite produced by flavin monooxygenase 3 (FMO3) in the liver from trimethylamine (TMA), increases in human and mouse blood as well as cerebrospinal fluid (CSF) with age, and TMAO treatment induces an increase in the aging phenotype, neuronal degeneration, and cognitive impairment in SAMP8, 3X Tg-AD mice, and HUVECs (Ke et al., 2018; Li et al., 2018; Govindarajulu et al., 2020; Brunt et al., 2021). Although the detailed mechanism remains unknown, it has been reported that it induces endoplasmic reticulum (ER) stress, C/EBP homologous protein (CHOP), and ROS to pass through the blood-brain barrier as well as the peripheral organs, causing neurodegeneration and affecting the CNS (Govindarajulu et al., 2020; Rosario et al., 2020). Even though the role of TMAO in PD pathogenesis is largely unknown, Chen et al. (2020) reported that increased plasma levels of TMAO in PD patients are associated with PD severity and progression.

This study devised a method to induce age-related phenotypes in midbrain organoids for aging and PD modeling. The effects of TMAO treatment were analyzed in terms of midbrain aging and PD pathogenesis. Hence, this study proposes a new research model to study late-onset brain diseases in human organoids and provides insight into how the gut microbiota can make the brain old.

Abbreviations: TMAO, trimethylamine *N*-oxide; PD, Parkinson's disease; AD, Alzheimer's disease; hPSCs, human pluripotent stem cells; ROS, reactive oxygen species; FMO3, flavin monooxygenase 3; TMA, trimethylamine; ER, endoplasmic Reticulum; CHOP, C/EBP homologous protein; CNS, central nervous system; hESC, human embryonic stem cell; CSF, cerebrospinal fluid; OD, optical density; SOX2, SRY-box transcription factor 2; FOXA2, forkhead box protein A2; TH, tyrosine hydroxylase; GFAP, glial fibrillary acidic protein; MAP2, microtubule-associated protein 2; *XBP1*, X-box binding protein 1; *GRP78*, glucose regulatory protein 78; CREB, cAMP response element-binding protein; pCREB, phosphorylated CREB; BDNF, brain-derived neurotrophic factor; CaMKII, calmodulin-dependent protein kinase II; PSD-95, post-synaptic density protein 95; *IL6*, interleukin 6; *IFNG*, interferon-γ; *TNFA*, tumor necrosis factor-α; RT-PCR, real-time PCR; HRP, horseradish peroxidase; TMB, tetramethylbensidine.

Materials and methods

Cell culture

H9 cells, a human embryonic stem cell line (WiCell, Madison, WI, United States) and human iPSCs (LRRK2^{G2019S} mutant PD-patient iPSC and gene-corrected iPSC) (Ha et al., 2020) were cultured in TeSR-E8 medium (STEMCELL Technologies, Vancouver, Canada) supplemented with 1% Pen/Strep (Gibco, Carlsbad, CA, United States) in matrigel (Corning, NY, United States) coated dishes. Cells were passaged using ReLeSR (STEMCELL Technologies, Vancouver, Canada), and the medium was changed on alternate days. This research with human embryonic stem cell and hiPSCs was approved by the Public Institutional Bioethics Committee designated by the Ministry and Welfare (MoHW) (Seoul, South Korea, IRB no. P01-201409-ES-01, P01-201802-31-001).

Generation of midbrain organoid

Human pluripotent stem cells maintained in the TeSR-E8 medium were dissociated into single cells with Accutase (MERTK), and 1.0 × 10^4 cells were seeded in an ultra-low attachment 96-well plate (S-bio, Hudson, NH, United States) for self-organization (Kwak et al., 2020). When the cell formed embryoid bodies, the media was replaced with EBM [DMEM/F12 (Gibco, Carlsbad, CA, United States) supplemented with 20% KSR (Gibco, Carlsbad, CA, United States), 50 μM Y27632 (Tocris, Bristol, United Kingdom), 3 μM CHIR99021 (Tocris), 1 μM IWP2 (Biogems, Westlake Village, CA, United States), 2 μM dorsomorphin (Sigma, St. Louis, MO, United States), 2 μM A83-01 (PeproTech, Rocky Hill, NJ, United States), 55 μM ß-mercaptoethanol (Gibco, Carlsbad, CA, United States), 3% FBS (Gibco, Carlsbad, CA, United States), 4 ng/ml bFGF (PeproTech), 1 μg/ml heparin (Sigma, St. Louis, MO, United States), 1% NEAA (Gibco, Carlsbad, CA, United States), 1% Pen/Strep, 1% GlutaMAX (Gibco, Carlsbad, CA, United States)], and after 24 h, it was replaced with BGM media (DMEM/F12: Neurobasal medium (1:1) supplemented with 1X N2 (Gibco, Carlsbad, CA, United States), 1X B27 w/o vitamin A (Gibco, Carlsbad, CA, United States), 3 μM CHIR99021, 1 μM IWP2, 2 μM dorsomorphin, 2 μM A83-01, 55 μM ß-mercaptoethanol, 1 μg/ml heparin, 1% NEAA, 1% Pen/Strep, and 1% GlutaMAX). After 2 days, FGF8 (PeproTech) and SAG (PeproTech) were added to the BGM media for mesencephalon patterning. The organoids were embedded in growth factor-reduced matrigel (Corning, NY, United States). Laminin (BD Science, Franklin Lakes, NJ, United States) and insulin (Thermo Scientific, Waltham, MA, United States) were added to the media, and CHIR99021, IWP2, dorsomorphin, and A83-01 were withdrawn. On day 9, the matrigel-embedded organoids were transferred to an ultra-low attachment 6-well plate and cultured on an orbital shaker at 60 rpm. The BMM medium [DMEM/F12: Neurobasal medium (1:1) supplemented with 1X N2, 1X B27 (Gibco, Carlsbad, CA, United States), 10 ng/ml BDNF (PeproTech), 10 ng/ml GDNF (PeproTech), 200 μM ascorbic acid (Sigma, St. Louis, MO, United States), 125 μM db-cAMP (Biogems), 55 μM ß-mercaptoethanol, 1 μg/ml heparin, 1% NEAA, 1% Pen/Strep, and 1% GlutaMAX] was changed on alternate days. After 4WM, 100 uM, or 1 mM of TMAO was continuously treated whenever the medium was changed once every 2 days.

RNA isolation and qPCR

The organoids were washed with 1XPBS and lysed using easy-BLUE^{TM}. Total RNA extraction kit (iNtRON Biotechnology, Seongnam, Republic of Korea). RNA was isolated using chloroform and isopropanol. The RNA pellet was washed with cold 70% EtOH, and the concentration was measured using Nanodrop. All 1,200 ng RNA was synthesized as cDNA with Superscript IV Reverse Transcriptase (Thermo Scientific, Waltham, MA, United States). The amount of gene expression was analyzed using Fast SYBR^{TM} Green PCR Master Mix (Applied Biosystems, Waltham, MA, United States) and primers for each marker. The expression level of each marker was normalized with *TBP*. The primer informations used in this study are provided in **Supplementary Table 1**.

Immunocytochemistry

The organoids were fixed in 4% paraformaldehyde at room temperature for 6 h and washed with 1XPBS. Next, organoids were embedded in OCT (Sakura, Tokyo, Japan) for frozen blocks, and frozen organoids were cut into 7 μm slices using a cryostat (LEICA CM1520). Samples were blocked with 3% of BSA and 0.02% of sodium azide in 0.025% TBS-T for 1 h after permeabilization with 0.3% of tritonX-100 (Sigma, St. Louis, MO, United States) in 0.025% TBS-T for 1 h, and target protein was stained with 1:100 diluted primary antibodies against SOX2 (Seven Hills Bioreagent, Cincinnati, OH, United States, #WRAB1236), FOXA2 (Seven Hills Bioreagent, Cincinnati, OH, United States, #WRAB1200), LMX1a (Sigma, St. Louis, MO, United States, #ab10533), Tuj1 (BioLegend, San Diego, CA, United States, #802001), TH (Sigma, St. Louis, MO, United States, #T1299), MAP2 (R&D, Minneapolis, Minnesota, United States, #MAB933), glial fibrillary acidic protein (GFAP) (DAKO, #Z0334), pERK1/2 (Cell signaling, #9106s), pCREB (Cell signaling, #9198S), p53 (Santa cruz, biotechnology, Dallas, TX, United States, #sc-126), Lamin A/C (Santa cruz, biotechnology, Dallas, TX, United States, #sc-376248), CaMKII (Novus biologicals, #NB110-96869), BNDF (Alomone labs, Jerusalem, Israel, #ANT-010), Tri-Me-K9 (abcam, Cambridge, United Kingdom, #ab8898), PSD-95 (Invitrogen, Waltham, MA, United States, #MA1-046), Synaptophysin (abcam, Cambridge,

United Kingdom, #ab32127), α-synuclein (Millipore, Billerica, MA, United States, #AB9850), pTau (Invitrogen, Waltham, MA, United States, MN1020), cleaved caspase 3 (Cell signaling, #9664s) at 4°C overnight. Slides were washed with 1X TBS-T and incubated with 1:400 diluted secondary antibodies (Alexa Fluor 488 goat anti-mouse IgG, Life Technologies, Alexa Fluor 555 donkey anti-rabbit, Life Technologies) with Hoechst (Thermo Scientific, Waltham, MA, United States) at room temperature for 1 h in deem light. The stained slides were observed under a fluorescence microscope (OLYMPUS, Shinjuku, Tokyo, Japan, U-TBI90) and a confocal microscope (ZEISS Oberkochen, Germany, LSM800).

Enzyme-linked immunosorbent assay

According to the manufacturer's instructions, we used a dopamine enzyme-linked immunosorbent assay (ELISA) kit to analyze the dopamine produced by midbrain organoids (Enzo, Basel, Swizerland, #ENZ-KIT188-0001). Briefly, conditioned media for each sample was added to each well in addition to the same volume of biotin detection antibody. The plate was set at 37°C for 45 min and washed three times. Next, the horseradish peroxidase (HRP) streptavidin conjugate working solution was added and set at 37°C for 30 min. The samples were then washed and incubated with the tetramethylbensidine (TMB) substrate for 15 min at 37°C in the deem light. The stop solution was immediately added, and the absorbance [optical density (OD)] was read using a microplate reader (SpectraMax, M3 Multi-Mode Microplate Reader) at 450 nm.

Fontana-Masson staining (melanin stain)

We visualized melanin in midbrain organoids to confirm the accumulation of neuromelanin using Fontana-Masson (Abcam, Cambridge, United Kingdom, #ab150669). The samples were washed in distilled water and set in an ammonical silver solution at 60°C for 45 min. Next, it was washed with distilled water and set in 0.2% gold chloride solution for 30 s at room temperature. The samples were then washed and set in 5% sodium thiosulfate solution at room temperature for 2 min. Finally, nuclei and cytoplasm were stained with the nuclear fast red solution for 5 min. Before mounting, samples were dehydrated in 100% EtOH.

Aggresome staining

The misfolded and aggregated protein were identified with the PROTEOSTAT Aggresome detection kit (Enzo, #ENZ-51035-K100). According to the manufacturer's instructions, samples were washed three times with 1XPBS for 10 min each. And then incubated with dual detection reagents (Dilute Aggresome Detection Reagent 1:2000 and Hoechst 33342 1:1000 in 1X Assay Buffer) for 30 min at room temperature in deem light. After washing with 1XPBS, the samples were mounted and observed under a confocal microscope.

Statistical analysis

All experiments were repeated three or more times and analyzed using Prism 6.0. Three to six organoid samples were analyzed for each group and used three sections per organoid. In each organoid section, three to five image was analyzed and image analysis proceeded with ImageJ software which quantified the fluorescence area of immunostaining in the same fluorescence intensity states and normalized with the fluorescence area of Hoechst. The non-parametric (Mann–Whitney) test was used for statistical analyses. Data values are presented as AVE ± SD. Statistical significance was set at $P < 0.05$, $P < 0.01$, $P < 0.001$, $P < 0.0001$ (*, **, ***, ****).

Results

Generation of midbrain organoid

We generated midbrain organoids from human embryonic stem cells (hESCs) to mimic the identity of the human midbrain, as previously reported by Kwak et al. (2020). Differentiation into midbrain organoids was performed by a guided neural-differentiation protocol (**Figures 1A,B**) and characterized by immunostaining for neuronal-specific markers (**Figures 1C,D**). At 2 weeks after maturation (2WM), we observed the expression of the neuronal stem cell marker SRY-box transcription factor 2 (SOX2), as well as the dopaminergic progenitor marker forkhead box protein A2 (FOXA2), which indicated midbrain-like specification of organoids. In addition, the expression of the dopaminergic neuronal marker tyrosine hydroxylase (TH) began to be detected in 2WM midbrain organoids (**Figure 1D**). In 4WM, the astrocyte marker GFAP was detected in organoids with microtubule-associated protein 2 (MAP2), TH and FOXA2, indicating glial differentiation in midbrain organoids (**Figures 1D,E**). In the magnified image, it was confirmed that GFAP-positive astrocytes were surround MAP2-positive neurons (**Figure 1F**).

The relative mRNA expression of cell type-specific genes was analyzed using RT-PCR for a quantitative comparison. The neural stem cell marker *SOX2* increased in 2WM organoids and then decreased in 4WM organoids, and the expression of the mature neuronal marker *MAP2* gradually increased with differentiation (**Figure 1G**). Dopaminergic neuronal markers (*PAX3*, *LMX1a*, and *TH*) maintained high

10.3389/fnagi.2022.925227

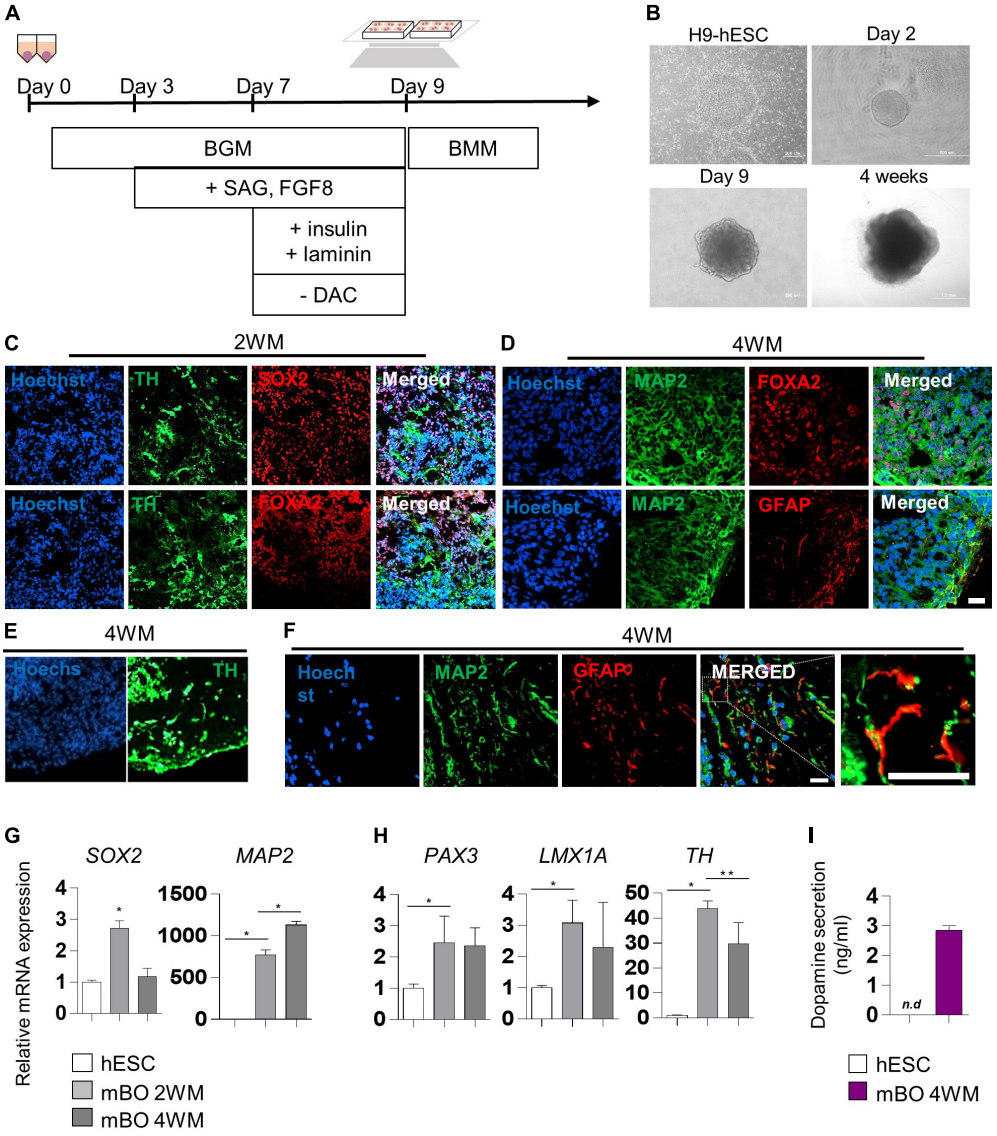


FIGURE 1

Characterization of midbrain organoids. **(A)** Schematic method of the midbrain organoids generation. **(B)** Representative phase-contrast images of human embryonic stem cell (hESC) and midbrain organoids at day 2, day 9, and 4 weeks of maturation. **(C–E)** Representative Immunostaining fluorescence (IF) Images showing the expression of the differentiation markers (*SOX2, FOXA2, TH, MAP2, and GFAP*) in midbrain organoids after 2, 4 weeks of maturation. Hoechst was used for counter nuclei staining. **(F)** High-resolution images of MAP2 and GFAP in midbrain organoids. **(G,H)** The relative mRNA expression level of differentiation markers (*SOX2, PAX3, LMX1a, TH, and MAP2*) of undifferentiated hESC and midbrain organoids (2WM, 4WM). **(I)** Quantification of dopamine secretion to conditioned media from undifferentiated hESC and midbrain organoids (4WMs). Data are AVE ± SD [$P < 0.05$(*), $P < 0.01$(**)]. (Scale bar = 20 μm).

at 4 weeks, but decreased slightly compared to 2 weeks. This is presumed to be due to the change in the relative portion of dopaminergic neurons in midbrain organoids with the increase of astrocytes and non-dopaminergic neurons (Figure 1H). Furthermore, the dopamine ELISA assay (Figure 1I) showed increased dopamine secretion (about 3 ng/ml) in the conditioned medium of single midbrain organoid cultures, demonstrating functional maturation of dopaminergic neurons in 4WM organoids.

Increase of cellular stress and aging features by trimethylamine *N*-oxide treatment

To evaluate the effects of TMAO on the neuropathological differentiation of the midbrain, we started treating 4WM organoids with TMAO at two concentrations (100 μM and 1 mM) for 26 weeks and analyzed the pathophysiological changes in organoids (Figure 2A). After 4 weeks of TMAO

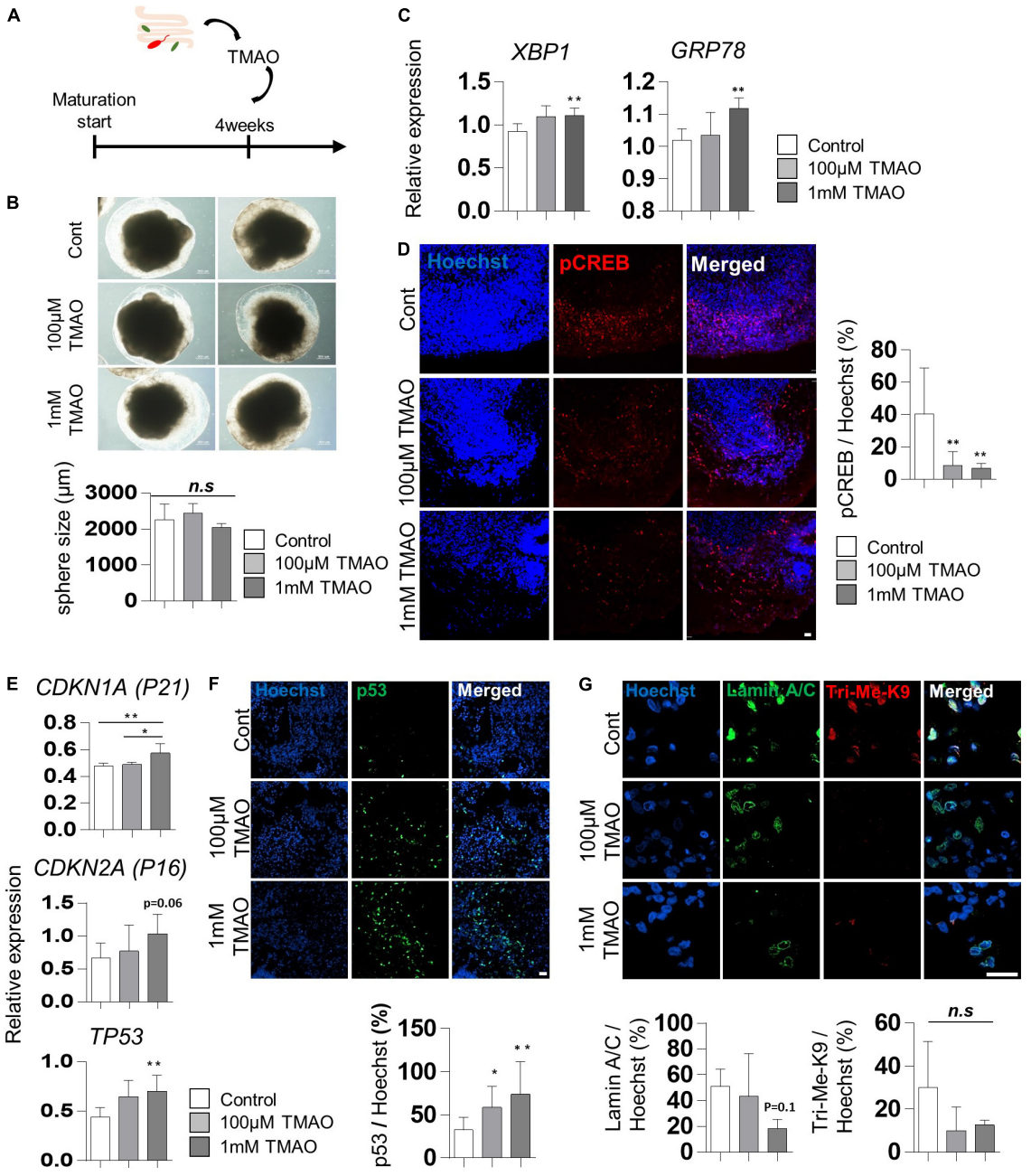


FIGURE 2
Aging-associated molecular phenotypes of trimethylamine *N*-oxide (TMAO)-treated midbrain organoids. **(A)** Schematics of TMAO treatment in midbrain organoid. **(B)** Representative phase-contrast images (upper) and quantification of organoid size (lower). **(C)** Relative mRNA expression levels of *XBP1* and *GRP78*. **(D)** Relative IF images (left) and quantification of phosphorylated cAMP response element-binding protein-positive cells in indicated groups (8WMs). Hoechst was used for counter nuclei staining. **(E)** The relative mRNA expression level of *CDKN1A, CDKN2A, and TP53*. **(F)** Representative IF images (upper) and quantification (lower) of the p53-expressing cells in the indicated organoid group. **(G)** Representative IF images (upper) and quantification (lower) of the Lamin A/C and Tri-Me-K9 in the indicated organoid group. Data are AVE ± SD [$P < 0.05$(*), $P < 0.01$(**)].

treatment (8WMs organoids), there was no significant difference in the diameter of the spheres and apoptotic cells between the control and TMAO treatment groups (**Figure 2B** and **Supplementary Figures 1A,B**).

Previous studies have reported that TMAO induces ER stress in the aged population and PD patients (Govindarajulu et al., 2020). In our experiments, the expression of the ER stress-related genes, X-box binding protein 1 (*XBP1*)

and glucose regulatory protein 78 (*GRP78*), was slightly increased in TMAO-treated organoids (**Figure 2C**), similar to that reported in pancreatic acinar cells (Yang and Zhang, 2021). In addition, phosphorylation of cAMP response element-binding protein (CREB), known to be inhibited by ER stress-mediated PERK activation (Kikuchi et al., 2016), was examined by immunostaining. The results showed a significant decrease of phosphorylated CREB (pCREB) by TMAO treatment in midbrain organoids (**Figure 2D**), implying the activation of ER stress by TMAO in midbrain organoids.

We analyzed changes in well-known senescence-associated aging markers in midbrain organoids to determine whether TMAO could induce aging-associated alterations in midbrain organoids. The expressions of *CDKN1A (P21), CDKN2A (P16)*, and *TP53* were analyzed by qPCR in 8WM organoids with or without treatment with TMAO for 4 weeks. The *TP53* expression was increased in organoids treated with TMAO at 100 μM and 1 mM concentrations compared to the control (**Figure 2E**). The expression of *CDKN1A* and *CDKN2A* also increased at the 1 mM TMAO treatment group (**Figure 2E**). In addition, immunostaining showed nuclear accumulation of p53 in TMAO-treated organoids (**Figure 2F**).

Moreover, we observed differences in the immunostaining results for lamin A/C and Tri-Me-K9, which are epigenetic aging markers reported to maintain the stability of the nuclear architecture and structure of chromatin (Scaffidi and Misteli, 2006). These proteins are defective with aging and lamin A/C and Tri-Me-K9 alteration, leading to telomere instability (Burla et al., 2016). Lamin A/C was decreased in the 1 mM TMAO treatment group compared to the control group. Tir-Me-K9 tended to decrease with TMAO (**Figure 2G**). These results demonstrate that TMAO reduces repressive epigenetic aging markers and decreases the structural stability of the nucleus and chromatin in midbrain organoids.

Weakened dopaminergic neuron protection by trimethylamine *N*-oxide treatment

Brain-derived neurotrophic factor regulates TH expression. It plays a protective role in dopaminergic neurons through CREB phosphorylation triggered primarily by MAPK/ERK1/2 and calmodulin-dependent protein kinase II (CaMKII) (Palasz et al., 2020). The decrease in phosphorylation of CREB by TMAO treatment was confirmed in **Figure 2D**. We further tested the intracellular signaling changes of BDNF supplemented in culture media to maintain the midbrain organoids. The phosphorylation of ERK1/2 in differentiated neuronal regions was dramatically reduced (**Figures 3A,C**), and CaMKII was decreased in TMAO-treated organoids. In contrast, BDNF expression was not significantly different (**Figures 3B,C**), implying impaired BDNF signaling.

Next, we compared the expression of TH, a marker of dopaminergic neurons, in midbrain organoids 4 weeks after TMAO treatment. Immunostaining showed that the number of TH-positive dopaminergic neurons was decreased in TMAO-treated organoids (**Figures 3D,F**). However, there was no significant difference in Tuj1, indicating that dopaminergic neurons were more susceptible to TMAO treatment than other neurons. The decrease in dopamine secretion from organoids was concentration-dependent following TMAO treatment (**Figure 3G**).

Synaptophysin and post-synaptic density protein 95 (PSD-95), pre- and post-synaptic proteins, respectively, were stained and quantified to confirm changes in synaptic proteins. PSD95 was decreased by TMAO treatment and Synaptophysin levels also slightly decrease with TMAO treatment, although this was insignificant (**Figures 3E,F** and **Supplementary Figure 2**), suggesting TMAO may affect synaptic dysfunction.

Furthermore, we detected neuromelanin in organoids synthesized from L-dopa and reportedly accumulates in the aged midbrain. Neuromelanin was produced in 8WM organoids, and its accumulation was significantly increased in the 1 mM TMAO treatment group than in the control group (**Figures 3H,I**). Altogether, the results show that TMAO-treated midbrain organoids could mimic parts of the cellular features of the aged midbrain, including loss of dopaminergic neurons, functional decline, and neuromelanin accumulation.

Astrocyte activation by trimethylamine *N*-oxide treatment

Given that astrocyte activation is one of the features of the aged brain, we confirmed the expression of the astrocyte activation marker, GFAP, in TMAO-treated midbrain organoids for 26 weeks. GFAP expression was enhanced by TMAO treatment, and astrocytes showed more activating morphology in the 1 mM TMAO treatment group, while TH-positive neurons were reduced (**Figure 4A**).

For quantitative analysis, *GFAP and hS100*β expression was examined by RT-PCR in 8WM organoids. As expected, *GFAP* expression was significantly increased in the TMAO treatment groups than in the control group. In addition, the expression of inflammatory cytokines, interleukin 6 (*IL6*) and interferon-γ (*IFNG*) was increased in the 1 mM treatment group compared to that in the control group. Still, there was no significant difference between the TMAO 100 μM treatment group and the control. Tumor necrosis factor-α (*TNFA*) levels were significantly increased by TMAO treatment (**Figure 4B**). These data show that TMAO might induce astrocyte-mediated inflammatory responses in midbrain organoids.

10.3389/fnagi.2022.925227

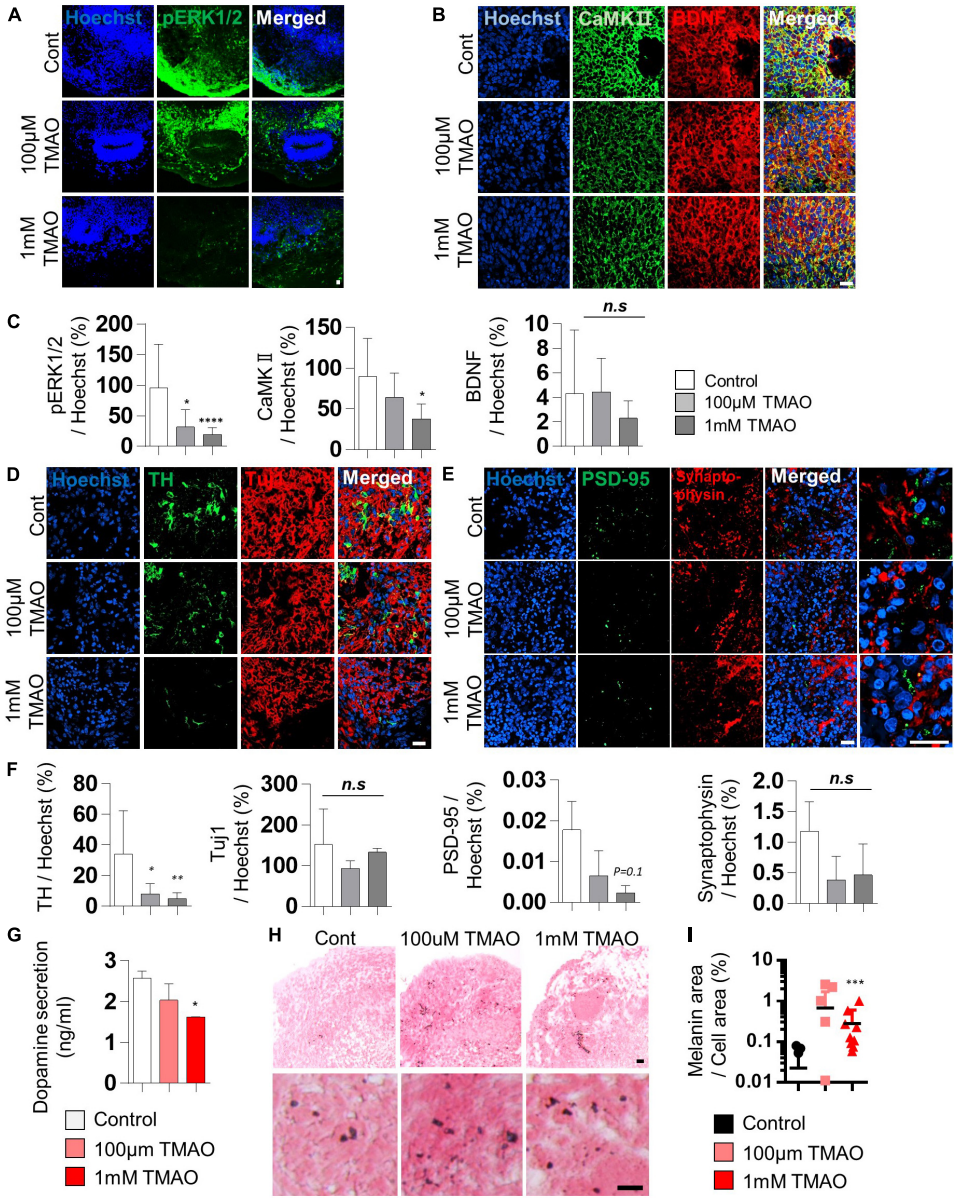


FIGURE 3
Neurodegeneration phenotypes of trimethylamine *N*-oxide (TMAO)-treated midbrain organoids. **(A,B)** Representative IF images showing the ERK phosphorylation **(A)** and CaMKII/BDNF **(B)** in the indicated organoid group. Hoechst was used for counter nuclei staining. **(C)** Quantifications of the ERK phosphorylation (left), CaMKII (middle), and BDNF (right) in the indicated organoid group. **(D,E)** Representative IF images showing the TH- and TUJ-1 positive neurons **(D)** and synaptic marker expression (PSD-95 and Synaptophysin) **(E)** in the indicated organoid group. Hoechst was used for counter nuclei staining. **(F)** Quantifications of the TH, TUJ-1 (left), PSD-95, and Synaptophysin (right) in the indicated organoid group. **(G)** Quantification of dopamine secretion to conditioned media from indicated organoids. **(H)** Representative Fontana-Masson staining images for neuromelanin at 8 weeks of maturation in midbrain organoids. **(I)** Quantifications of Fontana-Masson staining. Data are AVE ± SD [$P < 0.05$(*), $P < 0.01$(**), $P < 0.001$(***)]. (Scale bar = 20 μm).

Increase of abnormal protein aggregation by trimethylamine *N*-oxide treatment

To verify the association between TMAO and PD pathogenesis, we confirmed the phosphorylation of the α-synuclein protein at Ser-129 residues, a representative pathophysiological modification in PD (Anderson et al., 2006). Immunostaining of the TMAO-treated midbrain organoids for 26 weeks revealed reactive signals in the intracellular space of neurons (Figure 5A). Compared to the control group, the 1 mM TMAO treatment group showed an increase in phosphorylation

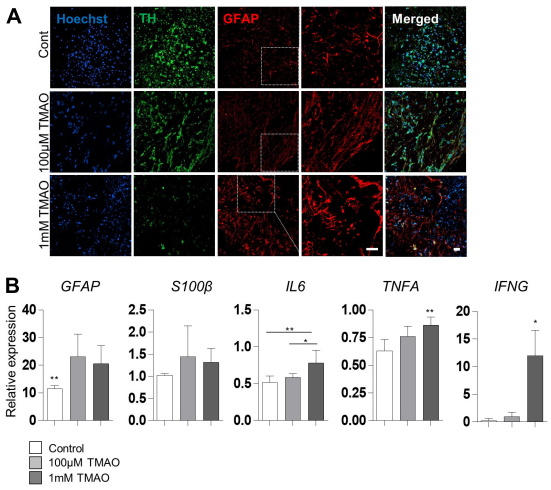


FIGURE 4
Astrocyte activations in trimethylamine *N*-oxide (TMAO)-treated midbrain organoids. **(A)** Representative IF images showing the TH-positive dopaminergic neurons and glial fibrillary acidic protein (GFAP)-positive astrocytes in indicated organoids. Hoechst was used for counter nuclei staining. **(B)** Relative mRNA expression level of GFAP and mRNA expression level of activation astrocytic markers, *GFAP, S100β, IL6, TNFA, and IFNA* in indicated organoids (8WMs). Data are AVE ± SD [$P < 0.05$(*), $P < 0.01$(**)]. (Scale bar = 20 μm).

at Ser-129, although not in 100 μM TMAO-treated organoids (**Figures 5A,B**).

Furthermore, phosphorylation of Tau at Ser202/Thr205 was increased by TMAO treatment. In the 1 mM TMAO treatment group, Tau accumulation was observed in the soma and the axons (**Figures 5C,D**). After confirming the increase in phosphorated α-synuclein protein and Tau, protein aggregation was confirmed by aggresome staining to determine whether abnormal aggregation of the protein increased. As a result, aggresome was significantly increased in 1 mM TMAO-treated organoids compared with a control group (**Figures 5E,F**).

Discussion

Since more than 90% of patients with neurodegenerative disorders, such as PD and AD, are sporadic cases, it is important to understand aging-associated alterations in the human brain and identify the relationship between age-related neurodegeneration and pathogenesis (Bekris et al., 2010; Slanzi et al., 2020). Herein, we investigated PD-associated predispositions of midbrain organoids by the microbiome metabolite TMAO, which increases with age. This study not only elucidates the role of TMAO in midbrain aging but also proposes a novel alternative method for studying late-onset degenerative brain disorders in hPSC-derived organoids.

Although an increase of TMAO has been reported in patients with PD, its role in midbrain aging and PD pathogenesis is largely unknown (Chen et al., 2020). In midbrain organoid models, we found that TMAO treatment can induce several features of the aged human brain, including increased cellular senescence and epigenetic aging markers, impaired neuronal function, glial cell activation, and inflammation (Mattson and Arumugam, 2018). Furthermore, we found that TMAO leads to phenotypes similar to the representative phenotype of PD, including loss of dopaminergic cells, increased phosphorylated alpha-synuclein, and phosphorylation of Tau. These results suggest that TMAO may play a crucial role in not only aging but also Parkinson's pathogenesis in the midbrain.

Increased senescent cells are a hallmark of aging, and the expression of p21 and p16 is closely related to cellular senescence (Bernardes de Jesus and Blasco, 2012; Capparelli et al., 2012; Lopez-Otin et al., 2013). In our organoid model, we observed an increased expression of p53, p21, and p16 following treatment with TMAO. In addition, epigenetics, known as an accurate cellular clock and epigenetic change, are strong indicators of biological aging (Galow and Peleg, 2022). Long-term TMAO treatment resulted in a tendency to decrease H3K9me3 in midbrain organoids, suggesting the possibility of epigenetic aging. Furthermore, we found defective nuclear lamina in the TMAO-treated organoids. As the a-type lamina (lamin A/C) assists in maintaining the nuclear structure and chromatin throughout the nucleus, it is associated with telomere distribution and function maintenance (Scaffidi and Misteli, 2006; Burla et al., 2016). These results suggest that the gut metabolite TMAO could induce cellular senescence-like molecular changes in midbrain organoids. Given that senescence cells accelerate tissue aging (Xu et al., 2018), TMAO-induced cellular senescence may play a role in midbrain organoid aging.

The decline in the number and function of dopaminergic neurons with normal aging is well-known, and excessive loss of dopaminergic neurons induces PD (Noda et al., 2020). We found that TMAO could induce organoid changes similar to the dysfunction of the aged midbrain. TMAO-treated organoids showed a progressive loss of dopaminergic neurons and reduced dopamine production. Moreover, decreased expression of PSD-95 in TMAO-treated organoids may imply synaptic dysfunction. On the other hand, the expression of the astrocyte marker GFAP was increased by TMAO. Increased GFAP expression is a common feature of reactive/activated astrocytes in the aged brain (Palmer and Ousman, 2018). In addition, we found increased levels of inflammatory cytokines, IL-6, TNFα, and IFNγ, in TMAO-treated organoids. Inflammation plays an important role in neurodegeneration and induction of aging-related mechanisms (Calabrese et al., 2018). The gut microbiome has also attracted attention as one of the factors involved in inflammaging (Franceschi et al., 2018). According to our results, TMAO may play a role in astrocyte-mediated

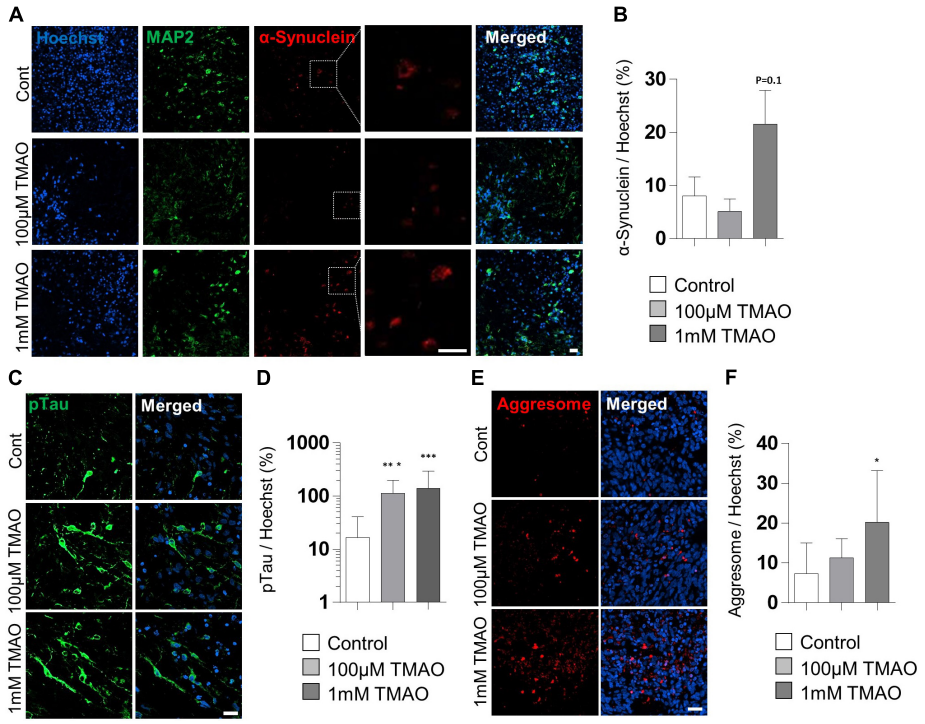


FIGURE 5
Parkinson's disease (PD)-associated phenotypes in trimethylamine *N*-oxide (TMAO)-treated midbrain organoids. **(A)** Representative IF images showing the microtubule-associated protein 2 (MAP2) and phosphorylated α-synuclein (ser129) in the indicated organoids (30WMs). **(B)** Quantification of phosphorylated α-synuclein (ser129) accumulation in neurons. **(C)** Representative IF images showing the phosphorylated-Tau (ser202/thr205) in the indicated organoids (30WMs). **(D)** Quantification of pTau accumulation in indicated organoids. **(E)** Representative IF images showing the aggresome in the indicated organoids (30WMs). **(F)** Quantification of aggresome in indicated organoids. Data are AVE ± SD [$P < 0.05$(*), $P < 0.01$(**), $P < 0.001$(***)]. (Scale bar = 20 μm).

inflammation response. However, since neurons, as well as glial cells, can secrete the pro-inflammatory cytokine IL6 (Sun et al., 2017), it is still possible that the increase in inflammatory cytokines by TMAO was caused by neurons, not activated astrocytes.

Moreover, we observed increased production of neuromelanin in TMAO-treated midbrain organoids. Neuromelanin synthesized from L-dopa is only found in primates, especially to be high in humans found in the substantia nigra and locus coeruleus (Fedorow et al., 2005). Considering the accumulation of neuromelanin in the elderly, it is presumed to be related to aging (Zucca et al., 2017). The increased accumulation of neuromelanin has been reported in patients with PD. Still, its role has rarely been reported, and it is also difficult to study due to the absence of neuromelanin in experimental models. Therefore, a TMAO-treated midbrain organoid model may be useful to study the role of neuromelanin in brain aging and PD pathogenesis.

Typical histological phenotypes of PD, such as Lewy bodies, were not observed in the TMAO-treated organoids. However, some predispositions to PD have been identified. As mentioned above, a decrease in the number and function of dopaminergic neurons was confirmed, and the phosphorylation of α-synuclein (Ser-129) was also increased in TMAO-treated organoids. In addition, an increase in pTau and aggresome was confirmed in our generated midbrain organoids. Although pTau has not been considered a pathophysiological feature of PD, it has recently been reported that pTau is increased in 50% of PD patients (Zhang et al., 2018). In particular, it is associated with tauopathy in sporadic PD. Also, an increase in pTau aggregation is a representative pathophysiological feature of AD. Therefore, the results of our study demonstrate the possibility of sporadic neurodegenerative disease modeling related to aging.

Recently, it was reported that TMAO induces CREB dephosphorylation *via* ER stress-mediated PERK phosphorylation, which leads to deficits in synaptic plasticity in an Alzheimer's mouse model (Chen et al., 2019; Govindarajulu et al., 2020; Yang and Zhang, 2021). In our midbrain organoid models, we also found a significant decrease in the pCREB levels in the TMAO treatment groups. In addition, the ER stress markers, *XBP1*, and *GRP78*, were increased by TMAO treatment. ER stress may play an important role in the TMAO-mediated midbrain organoid aging-like phenotype. Additionally, we newly demonstrated significant

dephosphorylation of ERK and reduction of CaMKII, which is known to play an important role in neuroprotection in a BDNF-dependent manner (Palasz et al., 2020). These results suggest that TMAO may have an important inhibitory effect on BDNF signaling, which can induce neurodegeneration in midbrain organoids. Further studies on the precise mechanisms and epidemiology of this condition are needed.

Although brain organoids provide a research model similar to the developing human brain, the aging signature remains a major obstacle to degenerative disease research. Currently, the most widely used method of inducing aging in organoids is long-term culture, but there are issues such as an increase in the study period, cost, labor and risk of contamination, and it also causes considerable damage in healthy-donor organoids, such as a massive increase in apoptosis. In this study, we suggest a novel method to accelerate the aging signature in stem cell organoids, especially midbrain organoid aging. We tried to recapitulate the natural aging in midbrain organoids by mimicking the increased gut metabolite TMAO in the elderly and diseased individuals. This method has the advantage of being able to simply induce the aging phenotypes without genetic modification, and it can recapitulate the primate-specific feature of midbrain aging such as the increase of neuromelanin. Thus, it will contribute to understanding the degenerative changes and pathogenesis in the human midbrain.

The application of TMAO-induced aging models in patient-derived iPSCs could lead to the advance of degenerative disease models. In our experiments, we generate midbrain organoids from a PD patient-derived iPSCs harboring LRRK2^{G2019S} mutation and isogenic WT iPSCs by gene correction (**Supplementary Figure 3A**). Unfortunately, because of the difference in the initial size growth and the degree of differentiation between mutant cells and correction cells, we could not thoroughly analyze the effect of TMAO treatment on PD pathogenesis (**Supplementary Figures 3B–F**). However, it was confirmed that TMAO treatment in these iPSC-derived organoids induced aging-related alteration of midbrain organoids, including reduced TH+ cells, decreased dopamine secretion, increased expression of p53 and p21, and decreased pERK (**Supplementary Figures 4A–F**). Our preliminary analysis with PD-patients iPSC has not shown an additional role of TMAO in PD pathogenesis, but the detailed relevance of PD and TMAO should be elucidated using more patient-based midbrain organoids. It is very interesting to see whether typical disease phenotypes such as Lewy body formation, which were difficult to reproduce in existing models, can be reproduced in TMAO-induced aging models.

It is also possible that TMAO is involved in midbrain aging but does not play an important role in PD pathogenesis. Considering that the role of TMAO in AD has been reported in mouse experimental animal models, it will be interesting to compare specific responses in various neuronal regions using regional specific organoids. By analyzing the TMAO response for each cell level, it is also possible to study the interaction between each cell type and the relationship between cellular senescence, neurodegeneration, and late-onset of disease. The absence of cell types such as blood vessels and microglia in current midbrain organoids will still act as a limitation in disease models using TMAO-induced aging.

In conclusion, we developed a TMAO-treated midbrain organoid model as a novel method to study midbrain aging in humans. The aging-like phenotypes of midbrain organoids were confirmed by increased expression of senescence markers, decreased expression tendency of repressive histone markers, neural degeneration, and neuromelanin accumulation. Although the TMAO-treated midbrain organoid model has limitations in reflecting all complex factors of natural aging, it will be useful for studying the mechanisms of brain challenges in the aged gut environment. Moreover, the combination of TMAO-induced aging organoids with genetic PD modeling, such as the LRRK2 mutation, would provide an opportunity to study the role of TMAO in the pathogenesis of PD, which could be an advanced human PD model.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding authors.

Ethics statement

This research with human embryonic stem cells and hiPSCs was approved by the Public Institutional Bioethics Committee designated by the Ministry and Welfare (MoHW) (Seoul, South Korea, IRB nos. P01-201409-ES-01, P01-201802-31-001).

Author contributions

YL, M-YS, and M-OL: conceptualization, manuscript writing, and review. YL, JK, and O-JH: validation and formal analysis. All authors have read and agreed with the published version of the manuscript.

Funding

This research was supported by grants from the National Research Foundation of Korea (NRF-2021R1C1C1006725, NRF-2018M3A9H3023077, and NRF-2021M3A9H3016046) funded by the Ministry of Science, ICT and Future Planning,

the Technology Innovation Program (No. 20009209) funded by the Ministry of Trade, Industry and Energy (MOTIE, South Korea), and the Korea Research Institute of Bioscience and Biotechnology (KRIBB) Research Initiative Program (1711134048, 1711134070, and 1711134076). The funders had no role in the study design, data collection or analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.925227/full#supplementary-material

References

Anderson, J. P., Walker, D. E., Goldstein, J. M., de Laat, R., Banducci, K., Caccavello, R. J., et al. (2006). Phosphorylation of Ser-129 is the dominant pathological modification of alpha-synuclein in familial and sporadic Lewy body disease. *J. Biol. Chem.* 281, 29739–29752. doi: 10.1074/jbc.M600933200

Azam, S., Haque, M. E., Balakrishnan, R., Kim, I. S., and Choi, D. K. (2021). The Ageing brain: molecular and cellular basis of neurodegeneration. *Front. Cell Dev. Biol.* 9:683459. doi: 10.3389/fcell.2021.683459

Bekris, L. M., Yu, C. E., Bird, T. D., and Tsuang, D. W. (2010). Genetics of Alzheimer disease. *J. Geriatr. Psychiatry Neurol.* 23, 213–227. doi: 10.1177/0891988710383571

Bernardes de Jesus, B., and Blasco, M. A. (2012). Assessing cell and organ senescence biomarkers. *Circ. Res.* 111, 97–109. doi: 10.1161/CIRCRESAHA.111.247866

Brunet, A. (2020). Old and new models for the study of human ageing. *Nat. Rev. Mol. Cell Biol.* 21, 491–493. doi: 10.1038/s41580-020-0266-4

Brunt, V. E., LaRocca, T. J., Bazzoni, A. E., Sapinsley, Z. J., Miyamoto-Ditmon, J., Gioscia-Ryan, R. A., et al. (2021). The gut microbiome-derived metabolite trimethylamine N-oxide modulates neuroinflammation and cognitive function with aging. *Geroscience* 43, 377–394. doi: 10.1007/s11357-020-00257-2

Burla, R., La Torre, M., and Saggio, I. (2016). Mammalian telomeres and their partnership with lamins. *Nucleus* 7, 187–202. doi: 10.1080/19491034.2016.1179409

Calabrese, V., Santoro, A., Monti, D., Crupi, R., Di Paola, R., Latteri, S., et al. (2018). Aging and Parkinson's disease: Inflammaging, neuroinflammation and biological remodeling as key factors in pathogenesis. *Free Radic. Biol. Med.* 115, 80–91. doi: 10.1016/j.freeradbiomed.2017.10.379

Capparelli, C., Chiavarina, B., Whitaker-Menezes, D., Pestell, T. G., Pestell, R. G., Hulit, J., et al. (2012). CDK inhibitors (p16/p19/p21) induce senescence and autophagy in cancer-associated fibroblasts, "fueling" tumor growth via paracrine interactions, without an increase in neo-angiogenesis. *Cell Cycle* 11, 3599–3610. doi: 10.4161/cc.21884

Chen, S. J., Kuo, C. H., Kuo, H. C., Chen, C. C., Wu, W. K., Liou, J. M., et al. (2020). The gut metabolite trimethylamine N-oxide is associated with Parkinson's disease severity and progression. *Mov. Disord.* 35, 2115–2116. doi: 10.1002/mds.28246

Chen, S., Henderson, A., Petriello, M. C., Romano, K. A., Gearing, M., Miao, J., et al. (2019). Trimethylamine N-Oxide binds and activates PERK to promote metabolic dysfunction. *Cell Metab.* 30, 1141–1151.e5. doi: 10.1016/j.cmet.2019.08.021

Fasano, A., Visanji, N. P., Liu, L. W., Lang, A. E., and Pfeiffer, R. F. (2015). Gastrointestinal dysfunction in Parkinson's disease. *Lancet Neurol.* 14, 625–639. doi: 10.1016/S1474-4422(15)00007-1

Fedorow, H., Tribl, F., Halliday, G., Gerlach, M., Riederer, P., and Double, K. L. (2005). Neuromelanin in human dopamine neurons: comparison with peripheral melanins and relevance to Parkinson's disease. *Prog. Neurobiol.* 75, 109–124. doi: 10.1016/j.pneurobio.2005.02.001

Franceschi, C., Garagnani, P., Parini, P., Giuliani, C., and Santoro, A. (2018). Inflammaging: a new immune-metabolic viewpoint for age-related diseases. *Nat. Rev. Endocrinol.* 14, 576–590. doi: 10.1038/s41574-018-0059-4

Galow, A. M., and Peleg, S. (2022). How to slow down the ticking clock: Age-associated epigenetic alterations and related interventions to extend life span. *Cells* 11:468. doi: 10.3390/cells11030468

Gandy, K. A. O., Zhang, J., Nagarkatti, P., and Nagarkatti, M. (2019). The role of gut microbiota in shaping the relapse-remitting and chronic-progressive forms of multiple sclerosis in mouse models. *Sci. Rep.* 9:6923. doi: 10.1038/s41598-019-43356-7

Ghosh, T. S., Shanahan, F., and O'Toole, P. W. (2022). The gut microbiome as a modulator of healthy ageing. *Nat. Rev. Gastroenterol. Hepatol.* 25, 1–20. doi: 10.1038/s41575-022-00605-x

Glowacki, R. W. P., and Martens, E. C. (2020). In sickness and health: Effects of gut microbial metabolites on human physiology. *PLoS Pathog.* 16:e1008370. doi: 10.1371/journal.ppat.1008370

Govindarajulu, M., Pinky, P. D., Steinke, I., Bloemer, J., Ramesh, S., Kariharan, T., et al. (2020). Gut Metabolite TMAO induces synaptic plasticity deficits by promoting endoplasmic reticulum stress. *Front. Mol. Neurosci.* 13:138. doi: 10.3389/fnmol.2020.00138

Ha, J., Kang, J. S., Lee, M., Baek, A., Kim, S., Chung, S. K., et al. (2020). Simplified brain organoids for rapid and robust modeling of brain disease. *Front. Cell Dev. Biol.* 8:594090. doi: 10.3389/fcell.2020.594090

Janeiro, M. H., Ramirez, M. J., Milagro, F. I., Martinez, J. A., and Solas, M. (2018). Implication of trimethylamine N-Oxide (TMAO) in disease: Potential biomarker or new therapeutic target. *Nutrients* 10:1398. doi: 10.3390/nu10101398

Ke, Y., Li, D., Zhao, M., Liu, C., Liu, J., Zeng, A., et al. (2018). Gut flora-dependent metabolite Trimethylamine-N-oxide accelerates endothelial cell senescence and vascular aging through oxidative stress. *Free Radic. Biol. Med.* 116, 88–100. doi: 10.1016/j.freeradbiomed.2018.01.007

Kikuchi, D., Tanimoto, K., and Nakayama, K. (2016). CREB is activated by ER stress and modulates the unfolded protein response by regulating the expression of IRE1alpha and PERK. *Biochem. Biophys. Res. Commun.* 469, 243–250. doi: 10.1016/j.bbrc.2015.11.113

Kim, H., Park, H. J., Choi, H., Chang, Y., Park, H., Shin, J., et al. (2019). Modeling G2019S-LRRK2 sporadic Parkinson's disease in 3D midbrain organoids. *Stem Cell Rep.* 12, 518–531. doi: 10.1016/j.stemcr.2019.01.020

Koo, B., Choi, B., Park, H., and Yoon, K. J. (2019). Past, present, and future of brain organoid technology. *Mol. Cells* 42, 617–627. doi: 10.14348/molcells.2019.0162

Kwak, T. H., Kang, J. H., Hali, S., Kim, J., Kim, K. P., Park, C., et al. (2020). Generation of homogeneous midbrain organoids with in vivo-like cellular

composition facilitates neurotoxin-based Parkinson's disease modeling. *Stem Cells* 38, 727–740. doi: 10.1002/stem.3163

Li, D., Ke, Y., Zhan, R., Liu, C., Zhao, M., Zeng, A., et al. (2018). Trimethylamine-N-oxide promotes brain aging and cognitive impairment in mice. *Aging Cell* 17:e12768. doi: 10.1111/acel.12768

Liu, C., Oikonomopoulos, A., Sayed, N., and Wu, J. C. (2018). Modeling human diseases with induced pluripotent stem cells: From 2D to 3D and beyond. *Development* 145:dev156166. doi: 10.1242/dev.156166

Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The hallmarks of aging. *Cell* 153, 1194–1217. doi: 10.1016/j.cell.2013.05.039

Mattson, M. P., and Arumugam, T. V. (2018). Hallmarks of brain aging: Adaptive and pathological modification by metabolic states. *Cell Metab.* 27, 1176–1199. doi: 10.1016/j.cmet.2018.05.011

Metta, V., Leta, V., Mrudula, K. R., Prashanth, L. K., Goyal, V., Borgohain, R., et al. (2022). Gastrointestinal dysfunction in Parkinson's disease: Molecular pathology and implications of gut microbiome, probiotics, and fecal microbiota transplantation. *J. Neurol.* 269, 1154–1163. doi: 10.1007/s00415-021-10567-w

Miller, J. D., Ganat, Y. M., Kishinevsky, S., Bowman, R. L., Liu, B., Tu, E. Y., et al. (2013). Human iPSC-based modeling of late-onset disease via progerin-induced aging. *Cell Stem Cell* 13, 691–705. doi: 10.1016/j.stem.2013.11.006

Mukherjee, A., Biswas, A., and Das, S. K. (2016). Gut dysfunction in Parkinson's disease. *World J. Gastroenterol.* 22, 5742–5752. doi: 10.3748/wjg.v22.i25.5742

Noda, S., Sato, S., Fukuda, T., Tada, N., and Hattori, N. (2020). Aging-related motor function and dopaminergic neuronal loss in C57BL/6 mice. *Mol. Brain* 13:46. doi: 10.1186/s13041-020-00585-6

Palasz, E., Wysocka, A., Gasiorowska, A., Chalimoniuk, M., Niewiadomski, W., and Niewiadomska, G. (2020). BDNF as a promising therapeutic agent in Parkinson's disease. *Int. J. Mol. Sci.* 21:1170. doi: 10.3390/ijms21031170

Palmer, A. L., and Ousman, S. S. (2018). Astrocytes and aging. *Front. Aging Neurosci.* 10:337. doi: 10.3389/fnagi.2018.00337

Pang, S. Y., Ho, P. W., Liu, H. F., Leung, C. T., Li, L., Chang, E. E. S., et al. (2019). The interplay of aging, genetics and environmental factors in the pathogenesis of Parkinson's disease. *Transl. Neurodegener.* 8:23. doi: 10.1186/s40035-019-0165-9

Potashkin, J. A., Blume, S. R., and Runkle, N. K. (2010). Limitations of animal models of Parkinson's disease. *Parkinsons Dis.* 2011:658083. doi: 10.4061/2011/658083

Rosario, D., Boren, J., Uhlen, M., Proctor, G., Aarsland, D., Mardinoglu, A., et al. (2020). Systems biology Approaches to understand the host-microbiome interactions in neurodegenerative diseases. *Front. Neurosci.* 14:716. doi: 10.3389/fnins.2020.00716

Scaffidi, P., and Misteli, T. (2006). Lamin A-dependent nuclear defects in human aging. *Science* 312, 1059–1063. doi: 10.1126/science.1127168

Schapira, A. H. V., Chaudhuri, K. R., and Jenner, P. (2017). Non-motor features of Parkinson disease. *Nat. Rev. Neurosci.* 18, 435–450. doi: 10.1038/nrn.2017.62

Simpson, D. J., Olova, N. N., and Chandra, T. (2021). Cellular reprogramming and epigenetic rejuvenation. *Clin. Epigenet.* 13:170. doi: 10.1186/s13148-021-01158-7

Slanzi, A., Iannoto, G., Rossi, B., Zenaro, E., and Constantin, G. (2020). In vitro models of neurodegenerative diseases. *Front. Cell Dev. Biol.* 8:328. doi: 10.3389/fcell.2020.00328

Sun, Q., Xie, N., Tang, B., Li, R., and Shen, Y. (2017). Alzheimer's disease: From genetic variants to the distinct pathological mechanisms. *Front. Mol. Neurosci.* 10:319. doi: 10.3389/fnmol.2017.00319

Wilmanski, T., Diener, C., Rappaport, N., Patwardhan, S., Wiedrick, J., Lapidus, J., et al. (2021). Gut microbiome pattern reflects healthy ageing and predicts survival in humans. *Nat. Metab.* 3, 274–286. doi: 10.1038/s42255-021-00348-0

Xu, M., Pirtskhalava, T., Farr, J. N., Weigand, B. M., Palmer, A. K., Weivoda, M. M., et al. (2018). Senolytics improve physical function and increase lifespan in old age. *Nat. Med.* 24, 1246–1256. doi: 10.1038/s41591-018-0092-9

Yang, G., and Zhang, X. (2021). TMAO promotes apoptosis and oxidative stress of pancreatic acinar cells by mediating IRE1alpha-XBP-1 pathway. *Saudi J. Gastroenterol.* 27, 361–369. doi: 10.4103/sjg.sjg_12_21

Zhang, X., Gao, F., Wang, D., Li, C., Fu, Y., He, W., et al. (2018). Tau pathology in Parkinson's disease. *Front. Neurol.* 9:809. doi: 10.3389/fneur.2018.00809

Zhu, S., Jiang, Y., Xu, K., Cui, M., Ye, W., Zhao, G., et al. (2020). The progress of gut microbiome research related to brain disorders. *J. Neuroinflammation* 17:25. doi: 10.1186/s12974-020-1705-z

Zucca, F. A., Segura-Aguilar, J., Ferrari, E., Munoz, P., Paris, I., Sulzer, D., et al. (2017). Interactions of iron, dopamine and neuromelanin pathways in brain aging and Parkinson's disease. *Prog. Neurobiol.* 155, 96–119. doi: 10.1016/j.pneurobio.2015.09.012

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 23 September 2022
DOI 10.3389/fnagi.2022.876832

OPEN ACCESS

EDITED BY
Yang Jiang,
University of Kentucky, United States

REVIEWED BY
Katya Numbers,
University of New South Wales, Australia
Ali Jawaid,
Nencki Institute of Experimental Biology (PAS), Poland

*CORRESPONDENCE
Lilla A. Porffy
lilla.a.porffy@kcl.ac.uk

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 15 February 2022
ACCEPTED 18 July 2022
PUBLISHED 23 September 2022

CITATION
Patchitt J, Porffy LA, Whomersley G, Szentgyorgyi T, Brett J, Mouchlianitis E, Mehta MA, Nottage JF and Shergill SS (2022) Alpha3/alpha2 power ratios relate to performance on a virtual reality shopping task in ageing adults. *Front. Aging Neurosci.* 14:876832. doi: 10.3389/fnagi.2022.876832

Alpha3/alpha2 power ratios relate to performance on a virtual reality shopping task in ageing adults

Joel Patchitt[1,2†], Lilla A. Porffy[1*†], Gabriella Whomersley[1], Timea Szentgyorgyi[1], Jack Brett[3], Elias Mouchlianitis[1,4], Mitul A. Mehta[1], Judith F. Nottage[1,5] and Sukhi S. Shergill[1,6,7]

[1]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom, [2]Trafford Centre for Medical Research, University of Sussex, Brighton, United Kingdom, [3]Faculty of Media and Communications, Bournemouth University, Poole, United Kingdom, [4]School of Psychology, University of East London, London, United Kingdom, [5]Department of Psychological Sciences, Birkbeck, University of London, London, United Kingdom, [6]Kent and Medway Medical School, Canterbury, United Kingdom, [7]Kent and Medway National Health Service and Social Care Partnership Trust, Kent, United Kingdom

Background: Aspects of cognitive function decline with age. This phenomenon is referred to as age-related cognitive decline (ARCD). Improving the understanding of these changes that occur as part of the ageing process can serve to enhance the detection of the more incapacitating neurodegenerative disorders such as Alzheimer's disease (AD). In this study, we employ novel methods to assess ARCD by exploring the utility of the alpha3/alpha2 electroencephalogram (EEG) power ratio – a marker of AD, and a novel virtual reality (VR) functional cognition task – VStore, in discriminating between young and ageing healthy adults.

Materials and methods: Twenty young individuals aged 20–30, and 20 older adults aged 60–70 took part in the study. Participants underwent resting-state EEG and completed VStore and the Cogstate Computerised Cognitive Battery. The difference in alpha3/alpha2 power ratios between the age groups was tested using *t*-test. In addition, the discriminatory accuracy of VStore and Cogstate were compared using logistic regression and overlying receiver operating characteristic (ROC) curves. Youden's J statistic was used to establish the optimal threshold for sensitivity and specificity and model performance was evaluated with the DeLong's test. Finally, alpha3/alpha2 power ratios were correlated with VStote and Cogstate performance.

Results: The difference in alpha3/alpha2 power ratios between age cohorts was not statistically significant. On the other hand, VStore discriminated between age groups with high sensitivity (94%) and specificity (95%) The Cogstate Pre-clinical Alzheimer's Battery achieved a sensitivity of 89% and specificity of 60%, and Cogstate Composite Score achieved a sensitivity of 83% and specificity of 85%. The differences between the discriminatory accuracy of VStore and Cogstate models were statistically significant. Finally, high alpha3/alpha2 power ratios correlated strongly with VStore (r = 0.73), the

Cogstate Pre-clinical Alzheimer's Battery (r = -0.67), and Cogstate Composite Score (r = -0.76).

Conclusion: While we did not find evidence that the alpha3/alpha2 power ratio is elevated in healthy ageing individuals compared to young individuals, we demonstrated that VStore can classify age cohorts with high accuracy, supporting its utility in the assessment of ARCD. In addition, we found preliminary evidence that elevated alpha3/alpha2 power ratio may be linked to lower cognitive performance.

KEYWORDS

virtual reality, cognition, neuropsychological testing, ageing, age-related cognitive decline, electroencephalography, alpha rhythm, cognitive markers

Introduction

Aspects of cognitive function decline with age (Murman, 2015), and these changes are not uniform across cognitive domains (Glisky, 2007). Crystallised intelligence (i.e., cumulative skills and acquired knowledge) improves with age and remains largely intact until late adulthood, whilst fluid intelligence (i.e., reasoning and problem solving) gradually diminishes (Murman, 2015). Evidence suggests that some fluid abilities (e.g., reasoning, spatial visualisation) begin to decline when healthy adults reach their late twenties and early thirties (Salthouse, 2009). This phenomenon is commonly referred to as age-related cognitive decline (ARCD). Improving the understanding of these cognitive changes that occur as part of the normal ageing process can serve to enhance the detection of the more incapacitating neurodegenerative conditions such as Alzheimer's disease (AD).

Despite increased knowledge of the brain mechanisms underlying neurodegenerative disorders, they are still primarily diagnosed using a combination of standardised neuropsychological assessments and subjective reports of cognitive decline (Stothart et al., 2021). This is because the assessment of biological markers, such as blood-based and image-based markers of beta-amyloid, tau, microglial activation, or glucose hypometabolism, often require more invasive procedures, and are high in cost and limited in availability (Frisoni et al., 2017). One technique, electroencephalography (EEG) provides a non-invasive and cost-effective approach to studying potential biomarkers. Indeed, a decline in cognition can be detected by rhythmic changes in frequency band power at rest (Babiloni et al., 2006). One promising EEG marker is the elevated relative power ratio of the high and lower alpha (α3/α2) band (Moretti et al., 2007). Increased α3/α2 power ratio was found in those patients with mild cognitive impairment (MCI) who converted to AD, but not in those who converted to non-AD dementias or did not convert during a 3-year retrospective follow-up study (Moretti et al., 2011a). In addition, increased α3/α2 power has been linked to hippocampal atrophy in patients with AD (Moretti et al., 2011b, 2012), and cognitive decline in patients with MCI (Moretti et al., 2013). The α3/α2 power ratio, however, has never been studied in a healthy ageing population; thus, it is unclear whether it is present in ARCD or specific to MCI and AD.

Existing neuropsychological measures also have several shortcomings. Standard assessments are prone to cultural bias and test-induced anxiety (Dorenkamp and Vik, 2018; Ng et al., 2018), lack ecological validity (Chaytor and Schmitter-Edgecombe, 2003), and insensitive to early stage cognitive dysfunction (Stothart et al., 2021). The most widely used screening tool for dementia, the Mini Mental State Examination (MMSE) (Folstein et al., 1975), is susceptible to both ceiling and floor effects; its outcome is influenced by education level, language, and cultural factors; and has no utility in the assessment of complex cognitive functions (Woodford and George, 2007). Newer computerised neuropsychological assessments may be more sensitive to early stage cognitive decline. The Cogstate Pre-clinical Alzheimer's Battery – measuring attention, processing speed, visual learning, and working memory – can reliably differentiate between healthy, MCI, and AD populations; with greater decline observed in those who carry the apolipoprotein E gene (APOE) ε4 allele, conferring increased for AD (Lim et al., 2012, 2015). However, the Cogstate Battery also has its limitations, with a recent prospective follow-up study showing that a single Cogstate assessment was not useful in predicting elevated beta-amyloid and tau levels in cognitively intact ageing individuals, while only achieving moderate diagnostic accuracy in predicting conversion from normal cognition to MCI with 77% sensitivity, 61% specificity (Stricker et al., 2020).

Cognitive assessments embedded in virtual reality (VR) may provide a solution to some of limitations associated with traditional neuropsychological testing. VR technology allows

for the development of three-dimensional (3D), interactive environments that resemble real-life situations, enhancing ecological validity and engagement (Bohil et al., 2011; Parsons, 2015). 3D spatial navigation is inherent to VR (Slater and Sanchez-Vives, 2016), placing additional cognitive load on the player similar to those elicited by real-life scenarios (Armougum et al., 2019). Indeed, a meta-analytic review found that VR tasks engage a greater range of cognitive domains compared to standard measures due to their increased complexity (Negut et al., 2016a). These attributes make VR-based tests sensitive to the early detection of dysfunctions in Instrumental Activities of Daily Living associated with cognitive decline (Romero-Ayuso et al., 2021). VR tasks have also demonstrated increased sensitivity in ageing healthy adults (Negut et al., 2016b). For example, age was the strongest predictor of performance on a recently developed VR test, ECO-VR, as opposed to education level or vocabulary skills (Oliveira et al., 2018). We recently reported that a novel VR shopping task, VStore, engages cognitive domains implicated in ARCD and AD, and is highly sensitive to predicting chronological age, both as a continuous and dichotomous outcome in healthy adults (Porffy et al., 2022a).

In this study, we employ novel methods to assess cognition in young and ageing cohorts of healthy individuals. First, we assess the utility of the α3/α2 power ratio EEG marker in ARCD. We expect that ageing participants have a higher α3/α2 ratio compared to young volunteers. Second, we aim to confirm our previous findings by testing whether VStore can discriminate between the two age cohorts and compare its discriminatory accuracy to Cogstate. Finally, to establish the relationship between these functional behavioural measures and EEG marker of cognition, we assess whether participants with a higher α3/α2 ratio have a lower performance on cognitive measures assessed by VStore and Cogstate.

Materials and methods

Participants

A total of 40 healthy volunteers aged 20–30 (n = 20) and 60–70 (n = 20) completed the study. Participants were recruited from South London *via* advertisements in local businesses, and on social media and local community websites. Participants were excluded if they had (1) a diagnosis of any Axis I disorder (DSM-5) (American Psychiatric Association, 2013), (2) alcohol and/or substance use disorder, (3) clinically significant motion sickness, (4) a neurological illness, (5) mobility issues, or (6) were pregnant. The sample size was determined based on the difference in overall VStore performance in our previous study (Porffy et al., 2022a) – see **Supplementary Material 1**. Hence, we decided to recruit 20 participant per age group. This sample size would give a sensitivity to detect an effect size of 0.91 or greater in α3/α2 ratio.

Behavioural measures

VStore

VStore is an immersive VR (IVR) assessment measuring functional cognition. The task takes approximately 30 min to complete including orientation (10 min), practice (10 min), and testing (10 min). Practice takes place in a virtual courtyard, where participants learn how to move around and manipulate objects. Once confident in the VR space, participants are teleported to a virtual minimarket where they complete the assessment.

At the start, 12 items are read out from a shopping list by a virtual avatar standing by the shop entrance (**Table 1**). The first task is to memorise and recall as many items from this list as possible. Following verbal recall, participants are presented with the shopping list and instructed to move around the minimarket and collect the items as fast and accurate as possible. Once all the items are collected, they are required to select and pay for them at a self-checkout machine, providing the exact amount. The task is completed by ordering a hot beverage from the coffee shop situated at the back of the store. The flowchart below summarises the steps required to complete VStore, its corresponding domains of cognition and outcome variables (**Figure 1**). Detailed information on VStore including development, feasibility, acceptability, tolerability, apparatus and software description, movement specification, and visual illustration has been published elsewhere (Porffy et al., 2022a,b).

Cogstate

Cogstate is a computerised cognitive battery designed to assess cognition across multiple domains. Cogstate is simple to use; therefore, it is deemed to be suitable for assessing older adults (Zygouris and Tsolaki, 2015). For the purposes of this study, eight tasks were selected measuring processing speed, attention, working memory, visual and verbal learning, executive functions, and paired associated learning (**Table 2**). Four of these – Detection, Identification, One Card Learning, One-back – have been validated as a Pre-clinical Alzheimer's Battery (Lim et al., 2012). We, therefore, used these tasks to calculate the Pre-clinical Alzheimer's Battery score by

TABLE 1 VStore shopping list.

1.	Cornflakes	7.	Colgate yoothpaste
2.	Tropicana orange juice	8.	Red apple
3.	Coca cola	9.	Raspberry jam
4.	Full fat milk	10.	Baked neans
5.	Tuna sandwich	11.	Orange
6.	Head and shoulders	12.	Brown bread

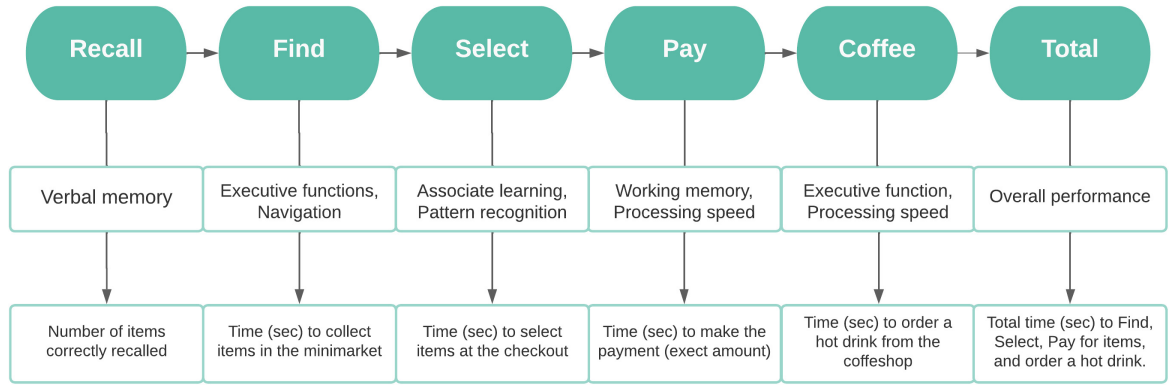


FIGURE 1
Flowchart depicting the steps required to complete VStore, its corresponding cognitive domains, and outcome variables. © Porffy et al. (2022b). Originally published in the Journal of Medical Internet Research (https://www.jmir.org), January 26, 2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0).

TABLE 2 List of cogstate tasks, corresponding cognitive domains, and main outcome measures.

Code	Task name	Cognitive domain	Outcome (metric)
DET	Detection	Processing speed	Reaction time (log 10 ms)
IDN	Identification	Attention	Reaction time (log 10 ms)
OCL	One Card Learning	Visual learning	Accuracy (arcsine proportion)
ONB	One-Back	Working memory	Reaction time (log 10 ms)
TWO	Two-Back	Working memory	Accuracy (arcsine proportion)
CPAL	Continuous Paired Associate Learning	Paired associate learning	Total number of errors
GMLT	Groton Maze Learning Task	Executive functions	Total number of errors
ISLT	International Shopping List	Verbal learning	Number of correct responses

standardising individual task scores and then averaging them within the whole sample including both young and ageing cohorts. We also generated a Composite Cogstate Score with the same procedure including all eight tasks.

Wechsler abbreviated scale of intelligence I

Intelligence quotient (IQ) was measured by the abbreviated version of the Wechsler Adult Intelligence Scale – first edition (Wechsler, 1999). The Wechsler Abbreviated Scale of Intelligence I (WASI-I) combines crystallised and fluid abilities to derive an age relative IQ score. The subtests used included the matrix reasoning and vocabulary tests.

Technological familiarity questionnaire

We developed a short questionnaire assessing technological familiarity (Porffy et al., 2022a,b). The Technological Familiarity Questionnaire (TFQ) includes 13 questions ascertaining the frequency, comfort, and ability in technology use on a 5-point scale. Higher scores indicate greater technological familiarity. In the present study, the TFQ total score was used to index of overall technological ability, while the item "how often do you use VR" was used to compare the frequency of VR use between age cohorts. The questionnaire's internal consistency was adequate ($\alpha = 0.83$, 95% CI = 0.82–0.89).

EEG resting-state recording

EEG recordings took place between 10:00–12:00 to control for time dependent electromagnetic fluctuations (Croce et al., 2018). Data was acquired using Compumedics Neuroscan SynAmps 64-channel amplifier. Participants were seated in a comfortable chair located in an unlit Faraday Cage with a computer monitor – displaying a fixation cross – 75 cm from their eyes. EEG measurements were taken using a 64 electrode Easycap, grounded to the AFz electrode. The channels were arranged according to the standard 10–10 system, except that some of the central electrodes (FC1, FC2, C1, C2, CP1, and CP2) were replaced by sub-temporal electrodes (FT9, FT10, P9, P10, POO9, and POO10) to better capture inferior temporal signals. All electrodes were referenced to the left mastoid. Vertical and horizontal eye movements were accounted for using an electrode above and below the right eye, and one on the outer canthus of the left eye. Impedance levels were reduced below 5 kΩ. Total recording time reached 4 min 30 s consisting of

eyes open and closed data. On average, 2 min 30 s of eyes closed data were obtained per recording. Participants were instructed through a headphone to open or close their eyes in 30 s intervals to reduce fatigue and the slowing of alpha waves (Ranlund et al., 2014). Data were acquired using a band pass filter of 0–1,000 Hz at a sampling rate of 5,000 Hz, and later digitised using a band-pass filter of 0.3–70 Hz at a sampling rate of 250 Hz as per previous studies in this frequency range (Moretti et al., 2012).

Procedures

Potential participants were pre-screened over the phone. Those eligible were invited for a single, 4-hour-long study visit. Informed consent was obtained at the start, followed by demographics and a brief mental and physical health history. EEG recordings were taken first, and the TFQ was completed during cap fitting. Upon acquisition of the EEG recording, participants were given a break to wash their hair and take lunch. Following the break, VStore and Cogstate were administered in a counterbalanced fashion, to minimise order effect. Finally, participants completed the WASI. Volunteers were compensated for their time and reimbursed for their travel and sustenance. Ethical approval was granted by the Psychiatry, Nursing and Midwifery Research Ethics Committee, King's College London (HR-18/19-11868).

Analysis

Behavioural data

Prior to data analysis, VStore outcomes measured in seconds were log transformed to stabilise variance. Outliers for all main VStore and Cogstate outcomes were defined using 2.5 standard deviations (SDs) above or below the sample mean as cut-off. Two participants, both in the 60–70 group, with outlier values on more than one VStore outcome were completely removed from all analyses. Additionally, a single extreme outlier value was removed from the Two-Back Cogstate task ($x = 0.51$, normal range = 0.84–1.77). For that participant, the Cogstate Composite Score was calculated based on average performance across the remaining 7 tasks. Furthermore, we found 3 more outliers among main outcome variables – one on VStore Pay, one on Cogstate Detection, and one on Cogstate Groton Maze Learning Task (**Supplementary Material 2**). For these, sensitivity analysis was carried out to establish whether they had impacted results. Descriptive statistics for VStore and Cogstate are presented in the supplementary (**Supplementary Material 3–5**).

Differences in demographic characteristics between age cohorts were tested using independent samples *t*-test and Chi-square. Group differences in cognitive performance – as measured by the Pre-Clinical Alzheimer's Battery, Cogstate Composite Score, and main VStore outcomes – were tested using independent samples *t*-test. The alpha level was adjusted using Bonferroni correction. Differences in VStore outcomes were considered significant at $\alpha = 0.009$, and at $\alpha = 0.025$ for Cogstate. Bootstrapped effect sizes (nboot = 1000) were calculated for each outcome using Hedges's *g*. To establish VStore's discriminatory accuracy, we built a logistic regression model using VStore Total Time as a predictor of group status (GroupStatus ~ VStore_T_Total). VStore Total Time was selected to represent a composite measure of overall performance. We repeated the procedure for the Cogstate batteries (GroupStatus ~ Cogstate_Alz; GroupStatus ~ Cogstate_Comp) and generated three overlying ROC curves. The analysis was repeated with the TFQ included to test its influence on results. Youden's J statistic was used to establish the optimal threshold for sensitivity and specificity and model performance was compared with the DeLong's test.

EEG data

Pre-processing and spectral analysis were carried out in EEGLAB and MATLAB (Delorme and Makeig, 2004; MATLAB, 2021). We removed ocular and power-line noise using independent component analysis. Next, we re-referenced to the average of left and right mastoids, and manually rejected any remaining artefacts and eyes open data. A Fast Fourier Transform-based power spectrum analysis calculated the spectral power density of the EEG rhythms with a frequency range of 2–40 Hz and a 0.5 Hz frequency resolution. The theta/alpha (θ/α) transition frequency (TF) and individual α frequency (IAF) were used as anchors to identify the subdivisions of the extended alpha spectrum (Klimesch, 1999). TF was computed as the minimum power within the extended α range (5–14 Hz) across all 64 channels. This identified where the θ/α frequencies intersect. IAF was computed as the highest averaged power, representing the "peak" within the same α frequency range. Using the TF and IAF, we isolated the absolute α3 and α2 band subdivisions. This was done by averaging the power values between the IAF to IAF + 2 hz for the α3 frequency band and calculating the same average between the middle point of the TF-IAF range to the IAF peak for the α2 frequency band (Moretti et al., 2007). Finally, relative spectral power was calculated for both for α3 and α2 band subdivisions by first, using the ratio between the absolute power for each frequency bin and mean spectra power from 2–45 hz, and then establishing the mean relative band powers from each frequency bin in that specific frequency band. The EEG dataset consisted of 35 participants. One participant was an outlier on both VStore and α3/α2 spectral power ratio outcomes (i.e., negative value for α3/α2), one person did not complete the EEG recording, and further three participants were excluded based on poor EEG data quality (i.e., no clear IAF peak). In total, 4 participants were removed from the 60–70 cohort and 1 participant was removed from the 20–30 cohort.

The remaining sample was assessed further for outliers. Values for two participants were found to be 2.5 SDs above

the group mean (Supplementary Material 2). Again, sensitivity analysis was applied to establish their influence on results. First, α3/α2 power ratios were compared between age cohorts using independent samples *t*-test. The effect size was calculated using Hedges' g. Second, Pearson's correlations were used to assess the associations between the α3/α2 ratios and VStore Total Time, and the Cogstate Pre-clinical Alzheimer's Battery and Composite Scores across the total sample. As per previous research (Moretti et al., 2012, 2013; Moretti, 2015b), we then grouped α3/α2 power ratios according to increasing tertiles. The boundaries were selected to create three equal groups classified as low (α3/α2 < 0.94, $n = 12$), mid ($0.94 \leq$ α3/α2 ≤ 1.02, $n = 11$), and high (α3/α2 ≥ 1.02, $n = 12$) α ratio. Chi-square test was used to assess the relationship between α3/α2 tertile groups and age cohort status. Finally, grouped α3/α2 ratios were separately correlated with VStote Total Time and the Cogstate batteries to establish how they relate to cognitive performance.

Results

Demographics

Demographic information for both cohorts is presented in Table 3. There was a significant difference in IQ and technological familiarity between groups. However, the frequency of past VR use was not higher in the 20–30 age cohort (16 never used VR) compared to the 60–70 age cohort (15 never used VR), $X^2_{(1)} = 0$, $p = 1$. The 60–70 age group only included people from a White ethnic background ($n = 18$, 100%), the 20–30 age group also included individuals from Asian ($n = 5$, 25%) and Mixed ($n = 3$, 15%) backgrounds.

Group differences in cognitive performance

The 60–70 age cohort were significantly slower on VStore outcomes Find, Select, Coffee, and Total Time compared to the 20–30 age cohort (Table 4). In addition, older volunteers achieved a lower score on the Cogstate Pre-clinical Alzheimer's Battery and had a lower Cogstate Composite Score. Results were unchanged following the removal of outliers. To further assess how performance on VStore relates to age in the 60–70 cohort, we run a Person's correlation between age and VStore Total Time (Supplementary Material 6).

TABLE 3 Sample characteristics.

	20–30	60–70	Statistics
N, count (%)	20 (53%)	18 (47%)	–
Age, mean (SD)	24.1 (3.0)	65.8 (2.4)	–
Gender, count (%)	10F (50%)	9F (50%)	$X^2_{(1)} = 0$, $p = 1$
IQ, mean (SD)	117.6 (8.2)	125.3 (7.1)	$t_{(35.9)} = -3.085$, $p = 0.004$
Education years, mean (SD)	16.1 (2.5)	16.2 (3.5)	$t_{(30.1)} = -0.172$, $p = 0.865$
Technological familiarity, mean (SD)	45.9 (6.2)	39.1 (9.0)	$t_{(29.8)} = 2.570$, $p = 0.012$

Group differences in α3/α2 ratio

The α3/α2 power ratios in the 20–30 age cohort (mean = 0.97, SD = 0.09) did not significantly differ from α3/α2 power ratios in the 60–70 cohort (mean = 1.07, SD = 0.26), $t_{(18.3)} = -1.514$, $p = 0.147$ (Hedges' $g = -0.518$, 95% = -1.14–0.12).

Age cohort classification

Figure 2 presents the discriminatory accuracy for VStore and Cogstate. VStore achieved a sensitivity of 94% and specificity of 95% at the optimal threshold of 0.47. The Cogstate Pre-clinical Alzheimer's Battery achieved a sensitivity of 89% and specificity of 60% at the optimal threshold of 0.37. Finally, the Cogstate Composite Score achieved a sensitivity of 83% and specificity of 85% at the optimal threshold of 0.38. The differences between the VStore model and both Cogstate models were statistically significant (Pre-Clinical Alzheimer's Battery: $z = 2.498$, *p*-value = 0.013; Cogstate Composite Score: $z = 2.020$, *p*-value = 0.043). Outliers did not have an impact on results, and the inclusion of the TFQ did not alter VStore findings. However, the TFQ decreased sensitivity (67%) and increased specificity (85%) for the Cogstate Pre-clinical Alzheimer's Battery, and increased sensitivity (94%) and decreased specificity (70%) for the Cogstate Composite Score. The AUC remained similar for all 3 models (< 1% difference).

Relationship between α3/α2 power ratios and cognitive performance

Across the complete sample, a significant positive association was found between α3/α2 power ratios and VStore Total Time; r = 0.35, $p = 0.04$. This finding was primarily driven by the moderate association between VStore Total Time and α3/α2 power ratios in the 60–70 cohort: $r = 0.36$, $p = 0.17$ (Figure 3). The relationship between VStore Total Time and α3/α2 power ratios; however, was no longer significant following the removal of outlier values; $r = 0.16$, $p = 0.38$. No correlation was found between α3/α2 power ratios and the Cogstate Composite Score; $r =$ -0.14,

TABLE 4 Group differences in cognitive performance between participants aged 20–30 vs. 60–70.

	20–30 Mean	60–70 Mean	*t*	*p*-value	Hedges' *g* (95% CI)
VStore recall	6.25	5.50	1.103	0.278	0.35 (-0.40–0.99)
VStore find	5.84	6.27	7.189	< 0.001	2.29 (1.38–3.03)
VStore select	4.63	5.10	4.974	< 0.001	1.58 (0.55–2.38)
VStore pay	2.90	3.07	1.280	0.209	0.41 (-1.05–0.38)
VStore coffee	3.34	3.78	3.902	< 0.001	1.23 (0.40–1.93)
VStore total	6.21	6.64	7.841	< 0.001	2.48 (1.50–3.25)
Cogstate alzheimer	0.42	−0.41	3.689	< 0.001	1.19 (0.50–1.80)
Cogstate composite	0.47	−0.45	4.496	< 0.001	1.44 (0.54–2.17)

Recall is presented as the number of correct responses. Find, Select, Pay, Coffee, and Total are presented in log transformed seconds.

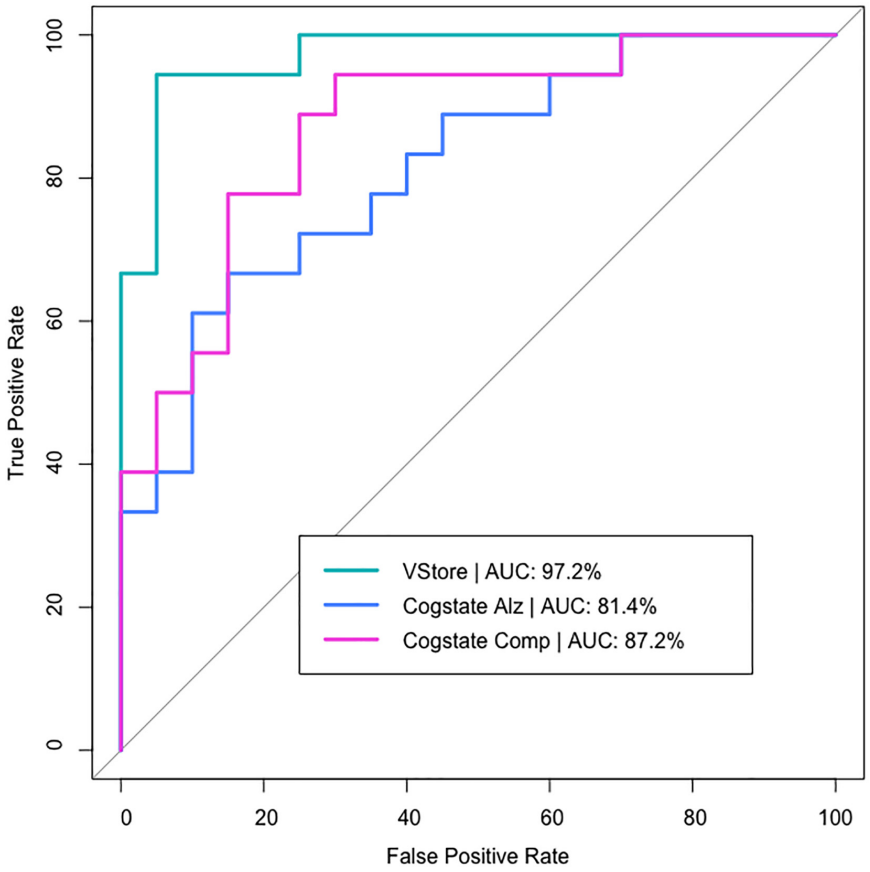


FIGURE 2
VStore and Cogstate models predicting age group belonging.

p = 0.43; or Cogstate Pre-clinical Alzheimer's Battery; r = -0.13, p = 0.45.

There was no relationship between α3/α2 ratio classification and age cohort status, $x^2_{(2)}$ = 1.237, p = 0.538 (Table 5). VStore Total Time showed a significant positive correlation with α3/α2 power ratios classified as high (r = 0.73, p = 0.01), but not with ratios in the mid (r = -0.32, p = 0.34), and low (r = -0.46, p = 0.13) ranges (Figure 4). A similar trend was observed with the Cogstate Pre-clinical Alzheimer's Battery, showing a significant association with high α3/α2 power ratios (r = -0.67, p = 0.02), but not with ratios in the mid (r = -0.12, p = 0.72), and low (r = 0.13, p = 0.69) ranges (Figure 5). The Cogstate Composite Score also showed a significant relationship with high α3/α2 power ratios (r = -0.76, p < 0.001), but not with mid (r = 0.08, p = 0.82), and low (r = 0.19, p = 0.56) ratios (Figure 6).

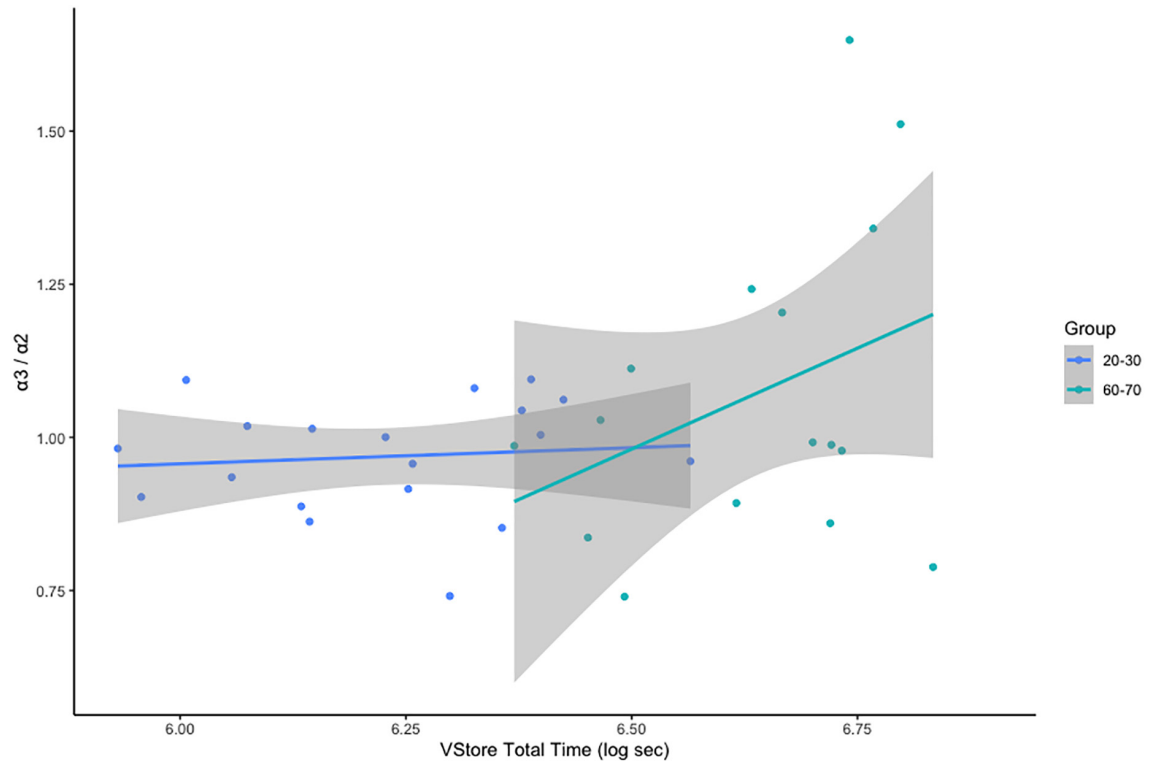


FIGURE 3
Correlations between VStore Total Time and α3/α2 ratios stratified by age group.

These results were somewhat altered following the removal of outliers, showing that VStore Total Time had the strongest association with the high α3/α2 power ratios (r = 0.74, p = 0.01), followed by the Cogstate Pre-clinical Alzheimer's Battery (r = -0.64, p = 0.04), and Cogstate Composite Score (r = -0.58, p = 0.06).

Discussion

To our knowledge, this is the first study to investigate α3/α2 resting state spectral power ratios in a healthy ageing cohort. Previous studies found elevated α3/α2 power ratios in individuals diagnosed with MCI and AD (Moretti et al., 2011b). We were unable to replicate these findings in healthy ageing adults. Previous, studies also showed that a high α3/α2 spectral power ratio is related to hippocampal atrophy (Moretti, 2015a). Given that participants in this study were high functioning, healthy adults, age-related changes may not have been pronounced enough to alter α3/α2 power. It is possible that the association between ageing and the high α3/α2 spectral power ratio is due to a degenerative process and is a marker of disease (Moretti et al., 2011a). Although to test this hypothesis, another sample of healthy participants and patients with MCI and AD would have to be tested on the EEG marker, hippocampal atrophy (Frisoni et al., 2008), and cognitive outcomes. Alternatively, negative findings could be due the lack of power to detect a difference in α3/α2 power ratios between age cohorts. In support, in healthy ageing individuals, evidence suggests that clusterin (CLU) gene polymorphisms, which increases AD risk, elevates α3 absolute power (Ponomareva et al., 2013). In addition, ageing has been linked to a decrease in α2 frequency power in multiple regions including limbic areas (Babiloni et al., 2006).

Receiver operating characteristic curve analysis revealed that VStore is highly sensitive and specific to the classification of the two age cohorts, more so than Cogstate, supporting its potential utility in the assessment of ARCD. These

TABLE 5 Frequency distributions of α3/2 power ratio tertiles by age group.

Cohorts	α3/2 tertiles			
	Low	Medium	High	Total
20–30	7	7	5	19
60–70	5	4	7	16
Total	12	11	12	35

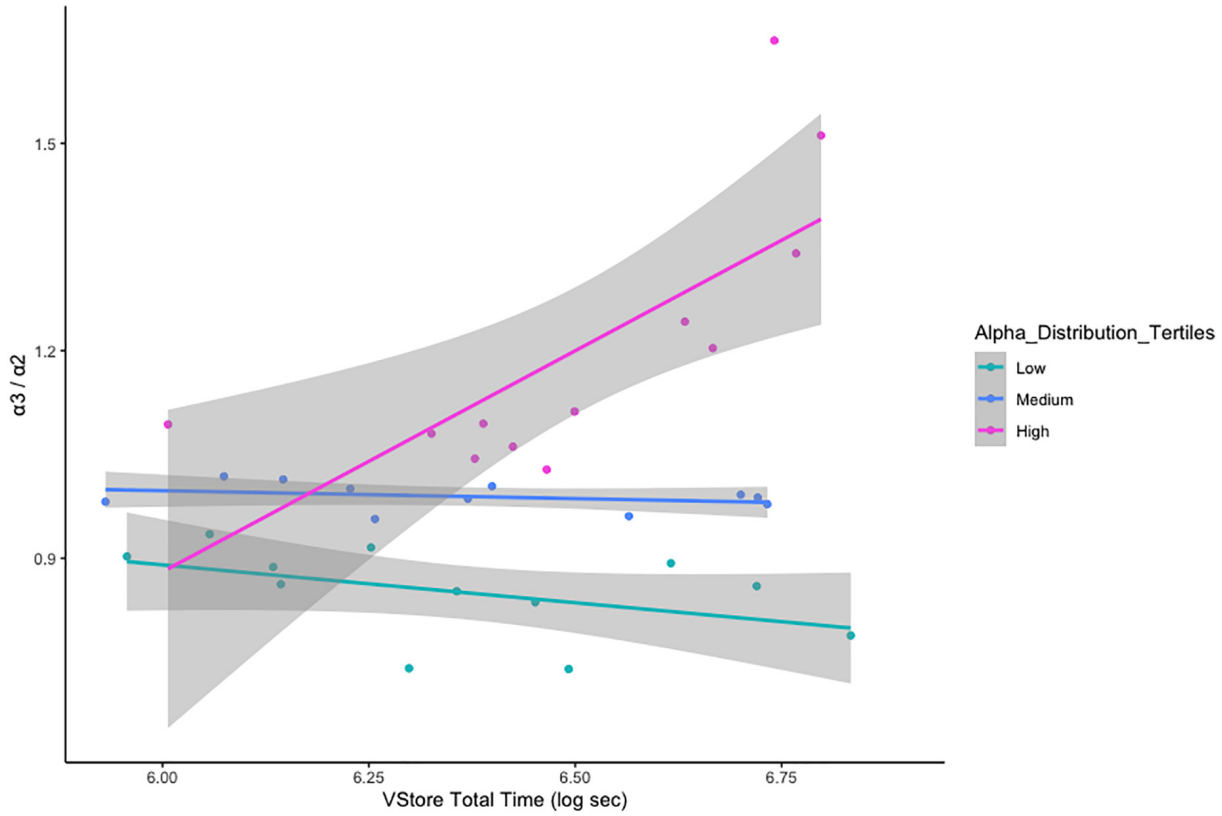


FIGURE 4
Correlations between VStore Total Time and α3/α2 ratios stratified by tertiles.

results confirm our previous findings (Porffy et al., 2022a), adding to a body of evidence that VR assessments may have an increased sensitivity in ageing healthy adults (Negut et al., 2016b). Indeed, the older cohort were slower to complete all VStore tasks, except making the payment. The VStore Pay outcome – engaging working memory, executive functions, verbal learning, and requiring fast processing speed – was amongst the weakest predictors of age as a continuous outcome in the previous study (Porffy et al., 2022a). This outcome, therefore, is likely to be less sensitive in ARCD. Interestingly, there was no marked difference between age cohorts in the number of VStore items correctly recalled. This is somewhat unexpected as verbal episodic memory tends to decline with age (Celsis, 2000). The lack of significant finding may be explained by the fact that the older cohort had a high verbal IQ on average, which may protect against memory decline (Boyle et al., 2021). Alternatively, immediate recall may not be as sensitive to cognitive decline as delayed recall (Gomar et al., 2011). Taken together, these findings suggest that assessments embedded in VR are sensitive to cognitive decline associated with ageing (Negut et al., 2016b; Oliveira et al., 2018; Porffy et al., 2022a), and also potentially provide valid concurrent measurement of everyday functioning (Romero-Ayuso et al., 2021); therefore, utilising this technology may enhance the early detection of subtle changes in cognition and related functional decline.

When we combined the EEG data from both cohorts, α3/α2 power ratios classified as high (≥ 1.02) showed a strong positive correlation with VStore Total Time, and strong negative associations with the Cogstate Pre-clinical Alzheimer's Battery and Composite Score. This suggests that elevated α3/α2 power ratios may be related to poorer cognitive performance. The relationship between α3/α2 power ratios and cognition has not been directly investigated in healthy individuals. However, in MCI those with a high α3/α2 power ratio, cognitive test performance correlates with cortical thickness (Moretti et al., 2013), providing indirect evidence that increased α3/α2 power may relate to a decline in cognition. Generally, changes in alpha power are linked to ageing as well as neurogenerative disorders, and lower alpha frequency correlates with poor memory and slow speed of processing (Klimesch, 1999). Alpha oscillation has also been linked to cognitive test performance on tasks probing attention, episodic memory, and executive functions in subjective and MCI (Babiloni et al., 2010). These domains are engaged during

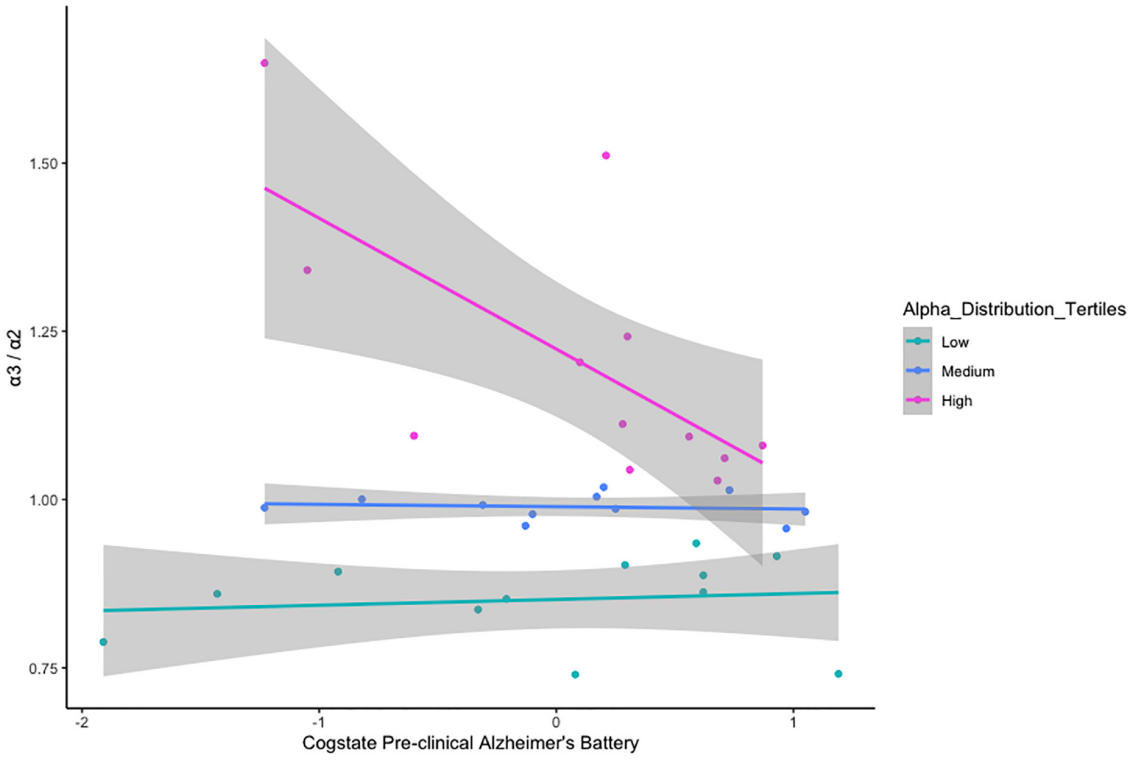


FIGURE 5
Correlations between the Pre-clinical Alzheimer's Battery and α3/α2 ratios stratified by tertiles.

VStore (Porffy et al., 2022a), which may explain the strong relationship between the EEG marker and VStore performance. Taken together, these findings suggest that VStore may be a useful screening tool for those whose α3/α2 power ratios may be elevated due hippocampal atrophy. However, this theory would have to tested.

There are some limitations to consider. VStore's high discriminatory accuracy may partly be due lower levels of technological proficiency in ageing adults (Olson et al., 2011). Indeed, ageing participants scored lower on technological familiarity. We tried to attenuate any potential confounding effects by providing sufficient practice time prior to assessment. In addition, we reran discriminatory models with the TFQ included. The addition of the TFQ did not have an impact on VStore results, suggesting that the high discriminatory accuracy was not simply due to differences in technological familiarity. This is in line with our previous findings indicating that the variance associated with technological familiarity is already captured in statistical models of VStore (Porffy et al., 2022a). In addition, we reported that past VR use was not more frequent in the young age group relative to the ageing cohort. Nonetheless, we cannot rule out that VStore, like any other digital assessment, may potentially underestimate the cognitive abilities of older adults. Furthermore, the sample included self-selecting ageing individuals and young students and professionals; therefore, it may not be completely representative of the general population. Opportunistic sampling tends to capture volunteers who are physically and mentally capable of attending and likely to exclude low functioning agers. Such problems are pervasive in this type of research (Murman, 2015), and could have affected EEG findings.

In conclusion, we did not find evidence that the α3/α2 spectral power ratio is elevated in healthy ageing individuals compared to young individuals. However, this may be due to the lack of sufficient statistical power, or the relatively high functioning of our volunteers. Further studies comparing cognitively intact young and ageing adults to patients with MCI and AD are needed to establish whether an increase in α3/α2 spectral power ratio is present in healthy ageing, supplemented with the measurement of hippocampal volumes. Despite these limitations, we confirmed previous findings showing that VStore can classify age cohorts with high accuracy, further supporting its utility in the assessment of ARCD. High-quality prospective studies are needed to establish whether VStore and other similar IVR assessments can detect early

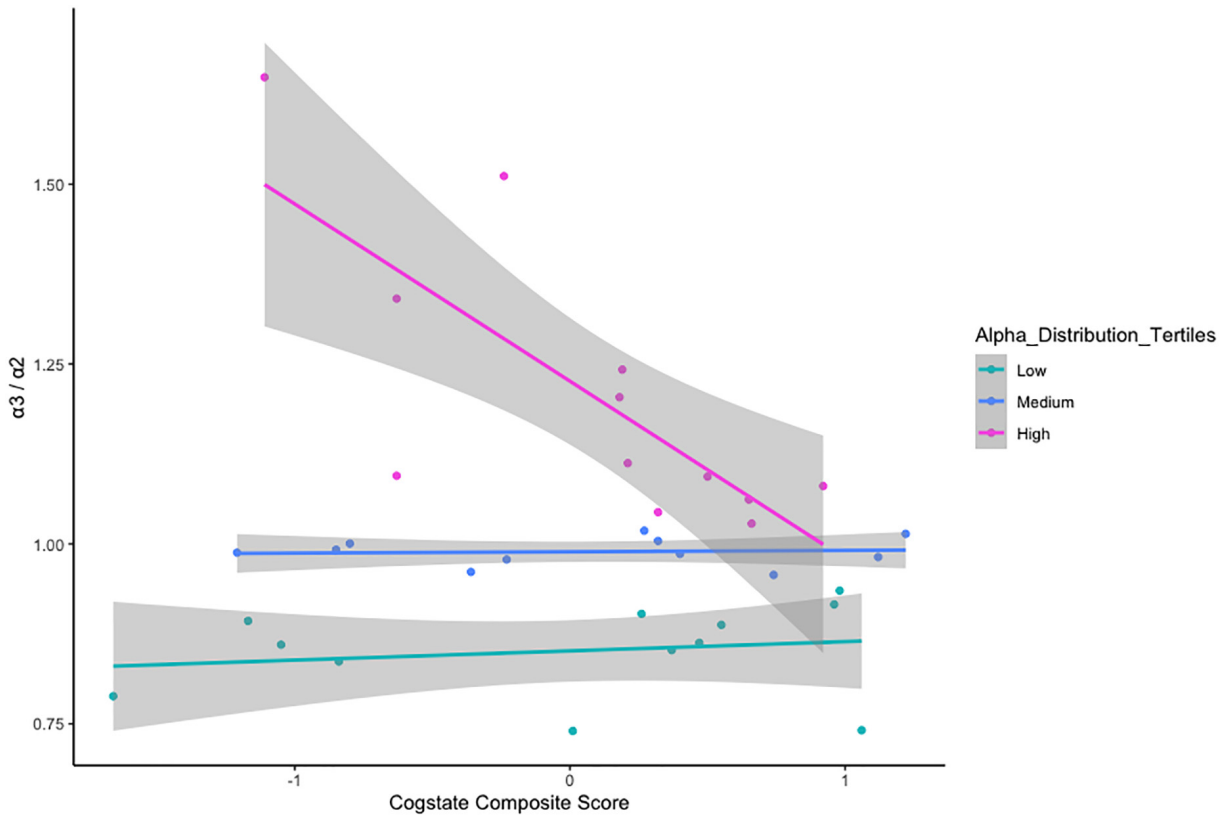


FIGURE 6
Correlations between the Cogstate Composite Score and α3/α2 ratios stratified by tertiles.

stage cognitive decline and provide a prognostic value in predicating transition to MCI and AD. Finally, we were able to establish that an increase in α3/α2 power ratio has a strong association to lower cognitive performance, suggesting that the marker may generally be useful in assessing cognitive decline.

Data availability statement

The datasets presented in this article are not readily available due to commercial restrictions. VStore is a propriety software. Requests to access the datasets should be directed to LP.

Ethics statement

The study involving human participants was reviewed and approved by the Psychiatry, Nursing and Midwifery Research Ethics Committee, King's College London (HR-18/19-11868). The participants provided their written informed consent to participate in this study.

Author contributions

JP: study and EEG protocol design, data collection and analysis, and contributed manuscript. LP: study design, data collection and analysis, and written manuscript. GW and TS: recruitment, data collection, and read and edited the manuscript. JB: VStore software development, and read and edited manuscript. EM: manuscript preparation and review. MM: advised on data analysis, and manuscript preparation and review. JN: EEG protocol design, advised on EEG data analysis, and read and edited manuscript. SS: study design and manuscript preparation and review. All authors contributed to the article and approved the submitted version.

Funding

This manuscript represents independent research funded by the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre (BRC) at South London and Maudsley NHS Foundation Trust and King's College London.

LP was supported by the United Kingdom Medical Research Council (MR/N013700/1) and King's College London member of the Medical Research Council Doctoral Training Partnership in Biomedical Sciences. The BRC Project Support Grant awarded to LP.

Acknowledgments

We thank volunteers for their participations in the study.

Conflict of interest

Authors SS and EM created VStore, with contributions from LP and technical development from VitaeVR Ltd., King's College London has licensed its rights in VStore to VitaeVR Ltd. Authors LP, EM, and SS are entitled to a share of any revenues King's College London may receive from commercialization of VStore by VitaeVR Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.876832/full#supplementary-material

References

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5§)*. Washington, DC: American Psychiatric Pub.

Armougum, A., Orriols, E., Gaston-Bellegarde, A., Marle, C. J.-L, and Piolino, P. (2019). Virtual reality: a new method to investigate cognitive load during navigation. *J. Environ. Psychol.* 65:101338. doi: 10.1016/j.jenvp.2019.101338

Babiloni, C., Binetti, G., Cassetta, E., Forno, G. D., Del Percio, C., Ferreri, F., et al. (2006). Sources of cortical rhythms change as a function of cognitive impairment in pathological aging: a multicenter study. *Clin. Neurophysiol.* 117, 252–268. doi: 10.1016/j.clinph.2005.09.019

Babiloni, C., Visser, P. J., Frisoni, G., De Deyn, P. P., Bresciani, L., Jelic, V., et al. (2010). Cortical sources of resting EEG rhythms in mild cognitive impairment and subjective memory complaint. *Neurobiol. Aging* 31, 1787–1798. doi: 10.1016/j.neurobiolaging.2008.09.020

Bohil, C. J., Alicea, B., and Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nat. Rev. Neurosci.* 12, 752–762. doi: 10.1038/nrn3122

Boyle, R., Knight, S. P., De Looze, C., Carey, D., Scarlett, S., Stern, Y., et al. (2021). Verbal intelligence is a more robust cross-sectional measure of cognitive reserve than level of education in healthy older adults. *Alzheimers Res. Ther.* 13:128. doi: 10.1186/s13195-021-00870-z

Celsis, P. (2000). Age-related cognitive decline, mild cognitive impairment or preclinical Alzheimer's disease? *Ann. Med.* 32, 6–14. doi: 10.3109/07853890008995904

Chaytor, N., and Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. *Neuropsychol. Rev.* 13, 181–197. doi: 10.1023/B:NERV.0000009483.91468.fb

Croce, P., Quercia, A., Costa, S., and Zappasodi, F. (2018). Circadian rhythms in fractal features of EEG signals. *Front. Physiol.* 9:1567. doi: 10.3389/fphys.2018.01567

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Dorenkamp, M. A., and Vik, P. (2018). Neuropsychological assessment anxiety: a systematic review. *Pract. Innov.* 3, 192–211. doi: 10.1037/pri0000073

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., et al. (2017). Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *Lancet Neurol.* 16, 661–676. doi: 10.1016/S1474-4422(17)30159-X

Frisoni, G. B., Ganzola, R., Canu, E., Rüb, U., Pizzini, F. B., Alessandrini, F., et al. (2008). Mapping local hippocampal changes in Alzheimer's disease and normal ageing with MRI at 3 Tesla. *Brain* 131, 3266–3276. doi: 10.1093/brain/awn280

Glisky, E. L. (2007). "Changes in cognitive function in human aging," in *Brain Aging, Models, Methods, and Mechanisms—Frontiers in Neuroscience*, ed. D. R. Riddle (Boca Raton, FL: CRC Press).

Gomar, J. J., Bobes-Bascaran, M. T., Conejero-Goldberg, C., Davies, P., and Goldberg, T. E. (2011). Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in Patients in the Alzheimer's disease neuroimaging initiative. *Arch. Gen. Psychiatry* 68, 961–969. doi: 10.1001/archgenpsychiatry.2011.96

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3

Lim, Y. Y., Ellis, K. A., Harrington, K., Ames, D., Martins, R. N., Masters, C. L., et al. (2012). Use of the CogState brief battery in the assessment of Alzheimer's disease related cognitive impairment in the Australian Imaging, Biomarkers and Lifestyle (AIBL) study. *J. Clin. Exp. Neuropsychol.* 34, 345–358. doi: 10.1080/13803395.2011.643227

Lim, Y. Y., Villemagne, V. L., Pietrzak, R. H., Ames, D., Ellis, K. A., Harrington, K., et al. (2015). APOE ε4 moderates amyloid-related memory decline in

preclinical Alzheimer's disease. *Neurobiol. Aging* 36, 1239–1244. doi: 10.1016/j.neurobiolaging.2014.12.008

MATLAB (2021). *MATLAB 9.10.0.16028886 (R2021a)*. Natick, MA: The MathWorks Inc.

Moretti, D. V. (2015b). Theta and alpha eeg frequency interplay in subjects with mild cognitive impairment: evidence from EEG, MRI and spect brain modifications. *Front. Aging Neurosci.* 7:31. doi: 10.3389/fnagi.2015.00031

Moretti, D. V. (2015a). Mild cognitive impairment: structural, metabolical, and neurophysiological evidence of a novel EEG biomarker. *Front. Neurol.* 6:152. doi: 10.3389/fneur.2015.00152

Moretti, D. V., Miniussi, C., Frisoni, G. B., Geroldi, C., Zanetti, O., Binetti, G., et al. (2007). Hippocampal atrophy and EEG markers in subjects with mild cognitive impairment. *Clin. Neurophysiol.* 118, 2716–2729. doi: 10.1016/j.clinph.2007.09.059

Moretti, D. V., Paternicò, D., Binetti, G., Zanetti, O., and Frisoni, G. B. (2013). EEG upper/low alpha frequency power ratio relates to temporo-parietal brain atrophy and memory performances in mild cognitive impairment. *Front. Aging Neurosci.* 5:63. doi: 10.3389/fnagi.2013.00063

Moretti, D. V., Prestia, A., Fracassi, C., Binetti, G., Zanetti, O., and Frisoni, G. B. (2012). Specific EEG changes associated with atrophy of hippocampus in subjects with mild cognitive impairment and Alzheimer's disease. *Int. J. Alzheimers Dis.* 2012:253153. doi: 10.1155/2012/253153

Moretti, D. V., Frisoni, G. B., Fracassi, C., Pievani, M., Geroldi, C., Binetti, G., et al. (2011a). MCI patients' EEGs show group differences between those who progress and those who do not progress to AD. *Neurobiol. Aging* 32, 563–571. doi: 10.1016/j.neurobiolaging.2009.04.003

Moretti, D. V., Prestia, A., Fracassi, C., Geroldi, C., Binetti, G., Rossini, P. M., et al. (2011b). Volumetric differences in mapped hippocampal regions correlate with increase of high alpha rhythm in Alzheimer's disease. *Int. J. Alzheimers Dis.* 2011:208218. doi: 10.4061/2011/208218

Murman, D. L. (2015). The impact of age on cognition. *Semin. Hear.* 36, 111–121. doi: 10.1055/s-0035-1555115

Negut, A., Matu, S. A., Sava, F. A., and David, D. (2016a). Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: a meta-analytic approach. *Comput. Hum. Behav.* 54, 414–424. doi: 10.1016/j.chb.2015.08.029

Negut, A., Matu, S. A., Sava, F. A., and David, D. (2016b). Virtual reality measures in neuropsychological assessment: a meta-analytic review. *Clin. Neuropsychol.* 30, 165–184. doi: 10.1080/13854046.2016.1144793

Ng, K. P., Chiew, H. J., Lim, L., Rosa-Neto, P., Kandiah, N., and Gauthier, S. (2018). The influence of language and culture on cognitive assessment tools in the diagnosis of early cognitive impairment and dementia. *Exp. Rev. Neurother.* 18, 859–869. doi: 10.1080/14737175.2018.1532792

Oliveira, C. R., Filho, B. J. P. L., Esteves, C. S., Rossi, T., Nunes, D. S., Lima, M. M. B. M. P., et al. (2018). Neuropsychological assessment of older adults with virtual reality: association of age, schooling, and general cognitive status. *Front. Psychol.* 9:1085. doi: 10.3389/fpsyg.2018.01085

Olson, K. E., O'Brien, M. A., Rogers, W. A., and Charness, N. (2011). Diffusion of technology: frequency of use for younger and older adults. *Ageing Int.* 36, 123–145. doi: 10.1007/s12126-010-9077-9

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Front. Hum. Neurosci.* 9:660. doi: 10.3389/fnhum.2015.00660

Ponomareva, N., Andreeva, T., Protasova, M., Shagam, L., Malina, D., Goltsov, A., et al. (2013). Age-dependent effect of Alzheimer's risk variant of CLU on EEG alpha rhythm in non-demented adults. *Front. Aging Neurosci.* 5:86. doi: 10.3389/fnagi.2013.00086

Porffy, L. A., Mehta, M. A., Patchitt, J., Boussebaa, C., Brett, J., D'Oliveira, T., et al. (2022a). A novel virtual reality assessment of functional cognition: validation study. *J. Med. Int. Res.* 24:e27641. doi: 10.2196/27641

Porffy, L. A., Mehta, M. A., and Shergill, S. S. (2022b). VStore: feasibility and acceptability of a novel virtual reality functional cognition task. *PsyArXiv [Preprint]* doi: 10.31234/osf.io/hg7wu

Ranlund, S., Nottage, J., Shaikh, M., Dutt, A., Constante, M., Walshe, M., et al. (2014). Resting EEG in psychosis and at-risk populations - a possible endophenotype? *Schizophr. Res.* 153, 96–102. doi: 10.1016/j.schres.2013.12.017

Romero-Ayuso, D., Castillero-Perea, Á, González, P., Navarro, E., Molina-Massó, J. P., Funes, M. J., et al. (2021). Assessment of cognitive instrumental activities of daily living: a systematic review. *Disabil. Rehabil.* 43, 1342–1358. doi: 10.1080/09638288.2019.1665720

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiol. Aging* 30, 507–514. doi: 10.1016/j.neurobiolaging.2008.09.023

Slater, M., and Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Front. Rob. AI* 3:1–47. doi: 10.3389/frobt.2016.00074

Stothart, G., Smith, L. J., Milton, A., and Coulthard, E. (2021). A passive and objective measure of recognition memory in Alzheimer's disease using fastball memory assessment. *Brain* 144, 2812–2825. doi: 10.1093/brain/awab154

Stricker, N. H., Lundt, E. S., Albertson, S. M., Machulda, M. M., Pudumjee, S. B., Kremers, W. K., et al. (2020). Diagnostic and prognostic accuracy of the cogstate brief battery and auditory verbal learning test in preclinical Alzheimer's disease and incident mild cognitive impairment: implications for defining subtle objective cognitive impairment. *J. Alzheimers Dis.* 76, 261–274. doi: 10.3233/JAD-200087

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Psychological Corporation.

Woodford, H. J., and George, J. (2007). Cognitive assessment in the elderly: a review of clinical methods. *Qjm* 100, 469–484. doi: 10.1093/qjmed/hcm051

Zygouris, S., and Tsolaki, M. (2015). Computerized cognitive testing for older adults: a review. *Am. J. Alzheimers Dis. Other Demen.* 30, 13–28. doi: 10.1177/1533317514522852

frontiers | Frontiers in Aging Neuroscience

TYPE Methods
PUBLISHED 18 January 2023
DOI 10.3389/fnagi.2022.952101

OPEN ACCESS

EDITED BY
Alessandro Martorana,
Tor Vergata University of Rome, Italy

REVIEWED BY
David W. Carley,
University of Illinois Chicago, United States
Ali Izadi,
University of California, Davis, United States

*CORRESPONDENCE
Christopher Daniel Morrone
✉ christopher.morrone@camh.ca
Wai Haung Yu
✉ waihaung.yu@camh.ca

SPECIALTY SECTION
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 24 May 2022
ACCEPTED 12 December 2022
PUBLISHED 18 January 2023

CITATION
Morrone CD, Tsang AA, Giorshev SM, Craig EE and Yu WH (2023) Concurrent behavioral and electrophysiological longitudinal recordings for *in vivo* assessment of aging.
Front. Aging Neurosci. 14:952101.
doi: 10.3389/fnagi.2022.952101

Concurrent behavioral and electrophysiological longitudinal recordings for *in vivo* assessment of aging

Christopher Daniel Morrone[1]*, Arielle A. Tsang[1,2], Sarah M. Giorshev[1,3], Emily E. Craig[1] and Wai Haung Yu[1,4,5]*

[1]Brain Health Imaging Centre, Centre for Addiction and Mental Health, Toronto, ON, Canada, [2]Department of Biological Sciences, University of Toronto Scarborough, Toronto, ON, Canada, [3]Department of Psychology, University of Toronto Scarborough, Toronto, ON, Canada, [4]Geriatric Mental Health Research Services, Centre for Addiction and Mental Health, Toronto, ON, Canada, [5]Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada

Electrophysiological and behavioral alterations, including sleep and cognitive impairments, are critical components of age-related decline and neurodegenerative diseases. In preclinical investigation, many refined techniques are employed to probe these phenotypes, but they are often conducted separately. Herein, we provide a protocol for one-time surgical implantation of EMG wires in the nuchal muscle and a skull-surface EEG headcap in mice, capable of 9-to-12-month recording longevity. All data acquisitions are wireless, making them compatible with simultaneous EEG recording coupled to multiple behavioral tasks, as we demonstrate with locomotion/sleep staging during home-cage video assessments, cognitive testing in the Barnes maze, and sleep disruption. Time-course EEG and EMG data can be accurately mapped to the behavioral phenotype and synchronized with neuronal frequencies for movement and the location to target in the Barnes maze. We discuss critical steps for optimizing headcap surgery and alternative approaches, including increasing the number of EEG channels or utilizing depth electrodes with the system. Combining electrophysiological and behavioral measurements in preclinical models of aging and neurodegeneration has great potential for improving mechanistic and therapeutic assessments and determining early markers of brain disorders.

KEYWORDS

electroencephalogram (EEG), mouse behavior, PhenoTyper, sleep, Barnes maze, wireless, aging

1. Introduction

Electrophysiological changes and disruptions in brain networking are associated with age-related cognitive decline and the development of neurodegenerative diseases (López-Sanz et al., 2019; McMackin et al., 2019; Saxena and Liebscher, 2020). Oscillatory signals captured by electroencephalography (EEG) correlate with behavioral deficits in patients with dementia and Alzheimer's disease (AD), and in animal models, and include alterations in neuronal frequency activity and power, cross-frequency coupling, and lagged-phase synchronization (Jeong, 2004; Canolty and Knight, 2010; Hamm et al., 2015; Hata et al., 2016). Impairments in EEG signals can even be observed in people with mild cognitive impairment, a precursor to dementias like Alzheimer's disease, making this a potentially useful tool for the early detection of dementia and neurodegeneration where subtle perturbations precede canonical disease neuropathology (Nakamura et al., 2018; Gaubert et al., 2019; van der Zande et al., 2020; Meghdadi et al., 2021). Despite the mounting evidence of neuronal network dysfunction affecting learning and memory in neurological and psychiatric disorders, preclinical techniques for longitudinal *in vivo* electrophysiological and behavioral assessments in freely moving animals are under utilized. This limits the ability to identify multimodal, synchronous changes such as electrophysiology and behavior affected by age and disease progression.

One such prominent contributor to, and the outcome of, neuronal network dysfunction is sleep abnormalities, which have long been known to occur (~50–60 years) in brain disorders (Ginzberg, 1955; Kupfer and Foster, 1972; Wang et al., 2015). Normal aging is associated with sleep disturbances in the daily sleep–wake cycle; however, the sleep loss associated with AD appears to be an exacerbation of these changes (Vitiello and Borson, 2001). Dementias are associated with hyperarousal states, including increased nighttime awakenings, decreased rapid eye movement (REM), and non-REM (NREM) sleep, in particular, delta-wave-dominant slow wave sleep or NREM stage 3 (Vitiello et al., 1991; Vitiello and Borson, 2001; Carter et al., 2010). These stages are normally critical for the cognitive and restorative benefits of sleep (Walker and Stickgold, 2004; Brown et al., 2012; Tononi and Cirelli, 2014). In fact, over 60% of patients with mild cognitive impairment or dementia report sleep disruptions, including a ~45–50% prevalence of insomnia (Guarnieri et al., 2012). The presence of sleep impairments confers a high risk for mild cognitive impairment and AD, notably in the development of AD biomarkers, 3.78× (Bubu et al., 2017), adding to the importance of investigating early EEG changes as it relates to behavioral changes. The correlative relationship between sleep and neurodegenerative disease may reciprocally drive disease progression, in which loss of sleep can exacerbate neurodegenerative processes, and the accumulation of proteinopathy can contribute to sleep impairments (Musiek and Holtzman, 2016; Minakawa et al., 2019). Critically, sleep disruption and hyperarousal are associated with aging and contribute to impaired cognitive performance in an array of neurodegenerative (i.e., AD, frontotemporal dementia, and Parkinson's disease) and psychiatric disorders (i.e., depression and schizophrenia) (Liu et al., 2004; Ju et al., 2014; Chahine et al., 2017; Fang et al., 2019; Waite et al., 2020; Wainberg et al., 2021).

Electroencephalography and behavioral testing have been frequently employed to research many disorders pre-clinically (Hamm et al., 2015; Pinnell et al., 2015, 2016; Xu et al., 2015; Kent et al., 2018; Medlej et al., 2019; Vorobyov et al., 2019; Buenrostro-Jáuregui et al., 2020; Garcia-Cortadella et al., 2021). Yet, in the study of aging and neurodegenerative disease, improving these techniques to preserve signal consistency over long periods of time is desirable. Furthermore, there is a need for these EEG recordings to be conducted in freely moving animals, facilitating a multitude of small and large arena behavioral tasks (i.e., Barnes maze) and providing a strong correlate of clinical assessments (Helfrich and Knight, 2019; Beppi et al., 2021). Herein, we employ a wireless EEG/EMG electrode headcap system to obtain high-quality longitudinal (up to 12 months post-surgery) EEG data from mice and describe the utility of combinatorial recordings during sleep, behavioral, and cognitive paradigms (**Figure 1**).

Abbreviations: AD, Alzheimer's disease; EEG, electroencephalography; EMG, electromyography; NREM, non-rapid eye movement; REM, rapid eye movement; SD, sleep disruption; TTL, transistor-transistor logic.

2. Materials and equipment

2.1. EEG headcap surgery

1. 2EEG/1EMG (stainless steel leads) headcaps (8201, Pinnacle Technology Inc.) – x1/mouse.
2. Anterior stainless steel electrode screws 0.10″ (Pinnacle Technology Inc.) – x2/mouse.
3. Posterior stainless steel electrode screws 0.12″ (Pinnacle Technology Inc.) – x2/mouse.
4. Flathead screwdriver, 1-mm head-size (available from Pinnacle Technology Inc.).
5. Stereotaxic frame, with nose cone for isoflurane and blunted ear bars.
6. Surgical tools (autoclaved or hot bead sterilized): scalpel handle and type #11 blade, fine forceps x3, curved forceps,

iris scissors, syringes (1 and 3 mL), and needles (23 and 30 g), hemostat, cotton swab x3–5/mouse.
7. Adhesives: cyanoacrylate Vetbond (various suppliers), 2-part epoxy (Pinnacle Technology Inc.), dental acrylic powder and jet liquid OR other dental cement (various suppliers); weigh boats and paintbrushes.
8. Animal care: Systane Gel lubricant eye gel (Alcon), fur clippers, ethanol (70%), Betadine scrub, Betadine solution, heat pad for mice, Dermachlor Rinse (Butler), sutures, recovery cage, alcohol wipes, and weighing scale; anesthetic: isoflurane; analgesic: Metacam, 5 mg/kg [dilute 1:10; the quantity injected is mouse weight (g) divided by 100]; local anesthetic: Bupivacaine, 1–2 mg/kg 0.125% (dilute 1:2; 0.1 mL injections along incision site); sterile gauze x3/mouse and sterile saline.

2.2. EEG recording equipment

1. Reusable potentiostat, Bluetooth USB dongle, battery caps (8274, Pinnacle Technology Inc.). We have tested up to 4 simultaneous 8274 recordings, with a potential for upward of 16 on one computer depending on Bluetooth traffic.
2. Zinc air batteries (Size 13, PR48, 1.45V; available from Pinnacle Technology Inc.).
3. Sirenia® Acquisition software (Pinnacle Technology Inc.).
4. Sirenia® Sleep Pro (Pinnacle Technology Inc.).

2.3. Behavioral equipment

1. PhenoTyper home cage and computer (Noldus).
2. Barnes maze (Maze Engineers): spatial cues, escape box, overhead light (LED; 5000K, 4000 lumen), camera, and computer (Noldus).
3. EthoVision XT15 software (Noldus).

3. Methods

3.1. Animals

All animal experiments were conducted in accordance with the ethical standards of the Canadian Council on Animal Care guidelines and approved by the Animal Care Committee of CAMH (Protocol #850). Mice were housed in a 12-h light:dark-cycle with *ad libitum* access to chow and water. For longevity analysis, 12 C57bl/6J mice (sex-balanced) underwent headcap surgery at 3.5 months of age. Of the 12, 6 were transgenic $App^{NL-F/NL-F}$ knock-in mice bred in-house (Saito et al., 2014) and 6 were non-transgenic mice (Jackson Laboratory). Ordered mice were allowed to habituate to the facility for 1.5 months prior to data collection. Four additional C57bl/6J female mice bred in-house were utilized for the optimization and validation of the EEG/EMG surgical technique and data analysis. To compare dark-cycle mobility without headcap and potentiostat, one additional non-transgenic mouse was utilized (female; bred in-house). Surgical implantation was successful in transgenic and non-transgenic, for both female and male mice; no sex differences were reported for frequency distribution or longevity of signal (refer to results section 4.1). **Table 1** summarizes mouse groupings utilized herein and the current surgical attrition rate (4.76%) for experiments following technical optimization and the completion of the present study. We report results with sex and genotype grouped together as our objective was to assess the longevity and viability of combining headcap recordings with behavior.

3.2. EEG headcap surgical procedure

Objective: 2EEG/1EMG headcap implantation allowing longitudinal recordings. **Note:** The surgical procedure was adapted from the information made available by Pinnacle Technologies Inc.

1. Ensure surgical equipment is sterilized and proper aseptic technique is followed.
2. Anesthetize the mouse in an isoflurane chamber (5% induction, 1% oxygen) and then transfer them to nosecone delivery.
3. Lower isoflurane to 3% and monitor the depth of anesthesia (i.e., toe pinch and breathing); isoflurane can now be adjusted downward in 0.25% intervals while maintaining the anesthetic plane. Once on the stereotaxic frame, maintain isoflurane at 1.25–2% while monitoring for the depth of anesthesia. Alternate anesthetics may be utilized but were not tested herein.
4. Weigh the mouse and then inject analgesic Metacam, 5 mg/kg (diluted 1:10), and saline (0.5 mL) subcutaneously.
5. Apply Systane Gel ointment on the eyes.
6. Inject local anesthetic Marcaine (bupivacaine), 1–2 mg/kg 0.125% (diluted 1:2); subcutaneously at the incision site (∼0.1–0.2 mL).
7. Using the clippers, shave away the fur on the top of the head.
8. Align the mouse in the stereotaxic frame, ensure the skull top is flat, and use a heat pad to maintain body temperature.
9. Clean the incision site with betadine scrub, then 70% ethanol, and then betadine solution.

10. Make a ∼1.5-cm incision using the scalpel blade along the skull midline, starting posterior to eyes until just past lambda.
11. Using a cotton swab, pull the skin to the sides to reveal the skull surface to avoid curling.
12. Dip another cotton swab in 70% ethanol to dry the skull surface (this ensures a strong adhesion of Vet Bond with potentiostat).
13. **CRITICAL STEP:** To align the placement of the headcap, use forceps to position the headcap with anterior screws to be embedded in the skull on top of the prefrontal cortex [(Bregma −2–2.5 mm AP, 1.5 ML (over both hemispheres)] and posterior screws to be anterior to lambda [(Bregma 3.5–4 mm AP, 1.5 ML (over both hemispheres)] above the retrosplenial and visual cortices (see Figure 1). A dye can be used to help landmark screw placement. **Note:** when the mice are small (20–25 g) it is advantageous to place the headcap as anterior as possible (before the skull becomes too narrow) to ensure that both anterior and posterior screws are above the cortex. If the posterior screws are further back where the cortex falls away and/or above the cerebellum, EEG activity on these channels will be too low in amplitude. **Note:** right hemisphere anterior and posterior screws confer signal for EEG2 and EEG1 channels, respectively; left hemisphere anterior and posterior screws confer signal for ground and EEG common (for reducing noise in the EEG1 and 2 channels), respectively.
14. Apply 1–2 drops of Vet Bond to the underside of the headcap once positioning is landmarked. Try to use as little as possible in order not to cover the screw holes, and leave it for 10 min for drying/curing.
15. **CRITICAL STEP:** Tap electrode screw holes with a 23 g needle, gently pushing down through the headcap holes to partially penetrate the skull surface. Rotate the needle lightly until there is mild resistance to create a pilot hole for the screw. Insert one screw at a time into the hole with fine forceps and use a 1-mm screwdriver to lower it into the hole. If it grips the skull (should feel minor resistance), then lower it until the screwhead is approximately halfway to the base of the headcap. If the screw is not gripping, continue to tap the hole with a needle and re-attempt the placement of the screw. Repeat for each screw until all 4 (2 anterior and 2 posterior) are in place.
16. Prepare epoxy fresh each time: mix 2 parts in a weigh boat with a paintbrush, stirring vigorously for 1 min. **Note:** Ensure epoxy is ready, this may require mixing or agitation of syringes first prior to application.
17. **CRITICAL STEP:** Apply a very small amount of epoxy to the screws using a fine gauge needle (e.g., 30 g needle on a 3-mL syringe as an applicator) on the outermost corner of the screw hole on the headcap so that it touches the threaded shank of the screw and connect the epoxy on the base of the headcap to the screwhead, ensuring a continuous seal. Repeat for each screw. Allow the epoxy to set for 15 min before applying dental acrylic. **Note:** do not let the epoxy from one screw contact another; a very minimal amount of epoxy is applied to not impede the signal.
18. Tighten screws so that the screwhead rests on the board base.
19. Placement of EMG wires in nuchal muscles: bend wires gently with fine forceps approximately half to two-thirds up so that the tip half is angled down (in line with how the neck is positioned in the stereotaxic frame) and laterally. With curved forceps make a pocket in the nuchal muscle and with fine forceps place the wire inside the pocket, ensuring that majority of the wire is within and covered by the muscle and not just under the skin. Repeat for the other wire.
20. **CRITICAL STEP:** Prepare dental acrylic/cement to apply to and insulate and protect screw electrodes: dip a small paintbrush in jet liquid and use a wet brush to pick up a small ball of acrylic powder, repeat dipping and picking up powder until the powder forms a gelatinous/viscous liquid ball for easier application. Apply to cover the screws and the base of the headcap, shaping it to continuously cover these parts and seal them to the skull.
    (a) Note: do NOT get acrylic on the top of the headcap (will block plugging-in of the potentiostat) or to the sides (will make it more difficult to grip the headcap when restraining the mice).
    (b) Note: covering the exposed edges of the EMG wires can help secure the wires and the headcap.
    (c) Note: over time (5+ months), the dental cement can wear away, so applying a full amount and ensuring the headcap is secured at the onset can help maintain the headcap placement. It is recommended to check every 3 months post-surgery and re-apply cement at least 1 week prior to collecting recordings.
    (d) Note: pull the skin away from cement to prevent it from fusing.
21. Suture the edges of the incision (typically only 1 suture is needed near the ears to close the skin above the back of the skull and the edges of EMG wires, check if an additional suture is necessary here and/or at the rostral end of the incision near the eyes).
22. Remove the mouse from the isoflurane and the stereotaxic frame and place them in a heated, clean cage for recovery.
23. During recovery: provide an additional injection of saline (0.5 mL) and Metacam analgesic [5 mg/kg (diluted 1:10)]. Use Dermachlor rinse or other topical antiseptics along the incision site and apply dropwise with minimal amounts.
24. Monitor recovery and post-op (additional analgesic as needed). Most mice recover within 30–45 min and exhibit normal behavior/are fully adapted

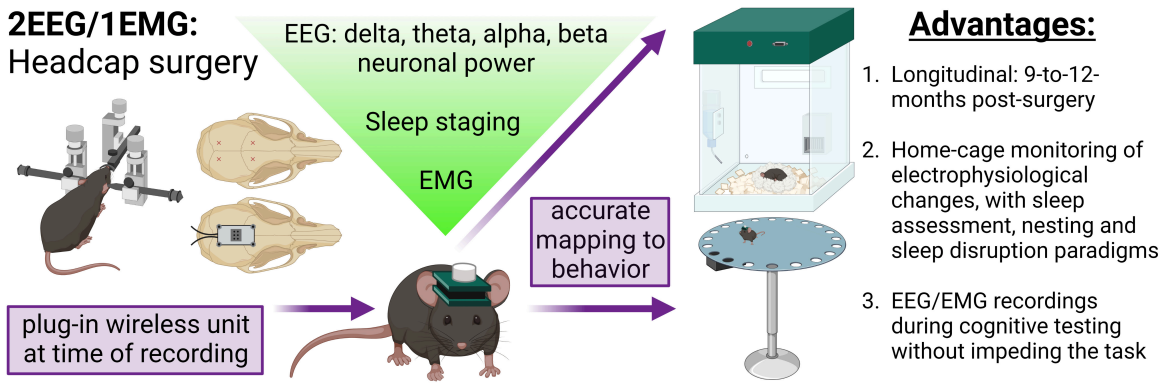


FIGURE 1
Concurrent EEG/EMG recordings with behavioral testing in mice. Schematic of the method employed. Mice underwent stereotaxic surgery to implant a headcap with two EEG channels and EMG wires in the nuchal muscle. EEG signal was derived from four stainless steel screw electrodes implanted at Bregma −2–2.5 mm AP, 1.5 ML, and Bregma 3.5–4 mm AP, 1.5 ML, as approximately indicated by the red × markings. A lightweight and small headcap was connected to the screws on the skull top and secured with dental cement. At the time of recording, a wireless potentiostat was plugged-in to acquire oscillatory signals. The EEG and EMG data can be utilized to assess alterations in neuronal frequency and sleep staging. Critically, the EEG/EMG system exhibits 9-to-12-month longevity and does not impede mice movement, allowing simultaneous electrophysiological and behavioral testing. In the present study, we demonstrate a combination of EEG/EMG recordings with home-cage locomotion, sleep assessments at baseline and after disruption, nesting, and, notably, during cognitive testing in the Barnes maze. We propose this system is feasible for concurrent recordings in a multitude of behavioral tasks beyond those investigated in the current study. Created with BioRender.com.

to the headcap at 24 h. Wait 5–7 days before handling and recording.

3.3. EEG/EMG recording and alignment with behavioral assessments

Objective: To set up recordings in time with behavioral assays and ensure the quality of the signal.

1. **CRITICAL STEP:** Pre-handling of mice is highly recommended to habituate the mouse and reduce stress when the experimenter must restrain it to plug-in potentiostats. Place the mouse on a cloth and keep a light grip on the base of the tail. After a few seconds, grip the sides of the headcap with a straight hemostat while maintaining a grip on the base of the tail. Practice letting go of the tail and keeping the mouse restrained *via* the grip on the headcap. This is when the potentiostat would be plugged in. Repeat for 3–5 days, 1–2× per day. **Note:** if the mouse is aggressive and/or repetitively jerking their head, release until agitation has lowered and try again.
2. On the day of recording, prepare for recording: connect USB dongles to the computer and open Sirenia® Acquisition software.
3. **CRITICAL STEP:** Remove the cover from the zinc batteries and let it sit for 5 min. Put the batteries in a potentiostat and cover it with the battery cap positioned with holes allowing airflow to the zinc batteries. Battery life allows upward of 72-h of recording at 1,024 samples/s. It is advantageous to set up a potentiostat Bluetooth connection early in case there is battery failure; ∼1-h before plugging-in the mouse.
4. Restrain mice as described in step 1 and plug them in the potentiostat.
5. Connect the potentiostat to the Bluetooth USB dongle receiver in Sirenia® Acquisition software. All data were sampled at 1,024 samples/s. EEG and EMG signals were acquired, digitized, and amplified at the potentiostat before being sent via Bluetooth to the USB receiver. Data were acquired at 100× gain and with 0.5 Hz (EEG1 and 2) and 10 Hz (EMG) high-pass filters, and a 500 Hz low-pass filter. EEG1 and 2 data were normalized to the EEG common electrode (left posterior screw), and the EMG signal was generated via the difference between the two wire leads.
6. **CRITICAL STEP:** Check the amplitude range [50–250 μV is an ideal range for EEG and EMG channels, depending on wake vs. sleep: EEG has higher amplitude during sleep, especially NREM, and EMG has higher amplitude during wake and variable], especially on first recording post-surgery to determine any issues with the electrode and wire placement.
7. Sirenia® Acquisition software encodes computer time within the file, allowing alignment with Noldus EthoVision XT software. EthoVision does not encode computer time but can be accurately controlled by the computer clock as set in the Trial Control Settings. For example, an experiment can be set to start accurately at a computer time

TABLE 1 Mouse groups utilized for optimization and analysis, and current attrition rates.

Mouse grouping	*N*	Sex	Analyses/Figures	Implant/Surgical concerns	Health/Aging concerns
Longevity (EEG/EMG headcap)	12 (50% $App^{NL-F/NL-F}$) transgenic, 50% non-transgenic).	6M 6F	8/12 mice: **Figures 3A–D** and sex differences analysis (4M and 4F); 7/12 mice: **Figures 3E–H**; Other four mice used for further testing (see below).	1 headcap dislodged @ 5-months post-surgery due to worn-away cement (see section "4.5 Pitfalls, limitations, troubleshooting, and alternatives").	1 malocclusion; 2 found dead; 1 sacrificed because skin grew under headcap.
Testing + young mice (4–5 months) (EEG/EMG headcap)	4 (extra from longevity; all non-transgenic)	2M 2F	4 mice: **Figures 4**, **6**, **7**; 3 mice: **Figure 5**; Supplementary data	None	See above; malocclusion endpoint before Barnes maze.
Testing + aged mice (9–10 months) (EEG/EMG headcap)	4	4F	4 mice: **Figures 4–6**; 2 mice: **Figure 7** Supplementary data	Improper EMG placement on 2 mice (EMG data omitted; see section "3.2 EEG headcap surgical procedure," step 19 for tips); 1 headcap dislodged @ 8-months post-surgery (after data collection) due to worn-away cement (see section "4.5 Pitfalls, limitations, troubleshooting, and alternatives"); 2 omitted from sleep disruption analysis due to battery issue.	None
Assessing mobility (no headcap)	1	1F	Supplementary data only	N/A	N/A
Other mice used in optimization of the current study	6	3M 3F	N/A	3 died in surgery; 1 sacrificed post-surgery (poor recovery).	N/A
Current attrition rates (for additional surgeries since completion of the current study)	**Total:** 3/42 mice (7.14%). **Surgery-related:** 2/42 mice (4.76%).	1M 2F	N/A	Surgery related endpoint. 1 died in surgery. 1 poor recovery.	Non-Surgery related endpoint. 1 weight loss.

The first five groupings document mice utilized in the present study to optimize this technique and assess its applications. The final row documents current attrition rates for experiments and surgeries conducted since the completion of this study (independent of the previous five groupings; no data shown from these mice), to provide an aid for experimental design.

of 12:00 in EthoVision, and by calculating the acquisition time, it can be linked to computer time as encoded by EEG/EMG data in Sirenia®.

8. Noldus PhenoTypers are home cages equipped with an overhead camera to monitor a variety of freely moving mouse behavior without the need for continued operator intervention. In the current application, this facilitated assessment of locomotion across a full circadian rhythm, collected in EthoVision XT. This provides a comprehensive sleep assessment in combination with EEG/EMG recordings. Other home-cage behaviors can be monitored alongside, including nesting as a measure of activities of daily living.

3.4. Barnes maze

Objective: To train mice to the location of the escape box (spatial learning) and assess their memory and neuronal activity during a probe recall trial (spatial memory). Seven mice were included in this analysis. Barnes maze was conducted similarly to what has been previously reported (Morrone et al., 2022; Xhima et al., 2022).

1. Around 1–2 weeks of mouse handling prior to Barnes maze assessment. For EEG/EMG mice, handle to familiarize the experimenter and mice, and utilize the technique described above to habituate the mice to plugging-in to the potentiostat.
2. A circular Barnes maze field (diameter: 92 cm; Maze Engineers; 20 holes) was utilized in a behavioral suite with an overhead camera connected to EthoVision XT15.
3. Testing Day 1: Mice were habituated to the Barnes maze and escape box, without spatial cues or the overhead light.
4. Overhead aversive light (brightness: 4,000 lumens; color range: 5,000 K) and spatial cues were included in all subsequent trials.
5. Testing Days 2–5: Mice were tested for learning in 2 trials per day (3-min trials with a 2-h inter-trial-interval). Once

the mouse enters the escape box, the light was turned off and the trial ended. The latency to find the escape box (s) and the number of errors were calculated in EthoVision to assess learning.

6. Testing Day 6: Following a 2-day break, mice were tested for 5 min in a probe trial in which the escape box was blocked. Thirty minutes prior to the probe, the potentiostat was plugged-in to allow habituation before being placed in the Barnes maze. As described above for the PhenoTypers, Sirenia® Acquisition and EthoVision software were run simultaneously. EEG/EMG data acquisition was started first (1 file per mouse) and later aligned to the clock start-time for the Barnes maze in the EthoVision raw data. The percentage of time spent searching in the correct quadrant was quantified in EthoVision.

3.5. Data analysis – Generating power and matching to the behavioral phenotype

Acquired EEG/EMG data was opened in Sirenia® Sleep Pro and exported by mouse and the time segment to be analyzed. A 50–60 Hz digital filter (band-stop filter) was applied to sampled raw EEG and EMG signals to remove potential electrical interference in the generated power data as well as for sleep staging. Power (μV^2/Hz) data was then generated by filtering out low- and high-band frequencies to capture the following neuronal waves: alpha (8–13 Hz), beta (13–30 Hz), delta (0.5–4 Hz), theta (5.5–8.5 Hz), and gamma (35–44 Hz); EMG: 50–150 Hz. These are the default frequency bins in Sirenia® Sleep Pro software which worked well for sleep staging in the present study, though there are options for adjusting these values (refer to section "4.5 Pitfalls, limitations, troubleshooting, and alternatives," point #7 for further discussion). The full range was analyzed for each channel as well in 0.5–500 Hz (EEG1 and 2) and 10–500 Hz (EMG). Power was generated in 10-s intervals for assessing longevity and matching neuronal frequency to locomotion, and in 2-s intervals for matching neuronal frequency to performance in the Barnes maze. Power data for longevity analysis is presented as an average of power (generated in 10-s intervals) within a 30-min period or split by wake and sleep state then averaged.

Power data was exported as a .tsv file and imported into Excel for further analyses and synchronization with behavioral readouts. Specifically, this involved alignment of data by computer clock time embedded in both datasets allowing direct comparison of electrophysiological and behavioral parameters at set times. **Note:** this requires both Sirenia® Acquisition and EthoVision XT to run simultaneously on the same computer.

The EEG2 neuronal power and EMG in 10-s epochs were averaged in 10-min bins to align to velocity (cm/s; velocity calculated in EthoVision as a 10-min average) for locomotor analyses; data were exported by mouse and the time segment to be analyzed: 7 p.m.–7 a.m. for locomotion dark cycle, 7 a.m.–7 p.m. for locomotion light-cycle.

For the Barnes maze, the entire recording was analyzed (5-min trial with ∼1 min of data preceding and following the test) and power was generated in 2-s intervals. EEG2 banded power was calculated as a percentage of full channel power and binned as target vs. non-target quadrants (based on location calculated in EthoVision), as well as 30-s baseline activity sampled from before and after the task. Data was further compared as power per band (alpha, beta, delta, theta, and EMG 50–150 Hz) between the target quadrant and non-target zones. Barnes maze power × location heat maps were generated in a representative mouse by first exporting raw positional data from EthoVision and averaging the 0.1-s *xy* data into 2-s intervals. The range of alpha and theta power was calculated to bin each 2-s interval as a proportion (0–10, 10–20, 20–30, 30–40, 40–50, 50–60, 60–70, 70–80, 80–90, and 90–100%) of total power. The location data was then graphed with the neuronal power indicated by size and color.

3.6. Generating representative data – EEG/EMG recordings to assess age-effects and response to PhenoTyper-mediated sleep disruption

Objective: To record EEG/EMG signal at baseline, during sleep disruption, and in a post-sleep disruption recovery period.

Young (4–5 months) and aged (9–10 months; refer to **Table 1**) mice underwent 24-h observation to capture a baseline circadian cycle of EEG/EMG recording. Mice were placed in PhenoTypers at 4 p.m. on the first day to allow habituation before recording onset at 7 p.m. (beginning of dark cycle). Video and EEG/EMG recordings occurred over 24 h (7 p.m. Day 1 – 7 p.m. Day 2), and mice were removed from the cage at 10 a.m. on the third day (after 42 h). **Note:** Both PhenoTyper and EEG systems allow recording up to 72 h, although only one full day was recorded herein. Nest images were captured at 18, 24, and 42 h and scored on a 1–5 scale (low-high complexity), as per Deacon (2006).

One week later, this paradigm was repeated with the addition of a 6-h sleep disruption (SD) period immediately before (12–6 p.m.). The top unit of PhenoTypers is additionally equipped with a white light and a speaker to elicit a tone (2,300 Hz, 80 dB). We used randomized white light and tone intervals as stimuli to reliably disrupt sleep, without direct handling of mice (Colavito et al., 2013). Specifically, every 30 s–3 min, the white light would turn on for a 20-s–1-min duration, and every 30 s–3 min, the tone would

play for 10–30 s, for the 6-h SD period. These settings are applied in the Trial Control Settings in EthoVision, under Action: Top Unit (hardware control); refer to **Supplementary Figure 2** for detailed settings. Notably, utilizing a home-cage environment with freely moving mice reduces stress, and PhenoTypers have built-in light and tone within each cage to increase reproducibility between animals and across studies.

Following SD, mice were immediately recorded by video and EEG/EMG for 24 h, and the nestlets scored, as in the baseline, facilitated a comparison of baseline and recovery after SD, as well as the effect of age. Nestlets were added part-way through the SD paradigm (4 p.m.). **Figure 2** depicts the timeline of these experiments.

3.7. Data analysis: Sleep staging

Epochs were scored as REM sleep, NREM sleep, or wake in 10-s intervals by cluster scoring in Sirenia® Sleep Pro; refer to **Supplementary Figure 1** for representative stages and cluster scoring, similar to previously described methods (Kent et al., 2018). Theta/delta ratio was utilized to distinguish REM (high theta/delta ratio and low EMG) and NREM (low theta/delta ratio, high delta power, and low EMG) from wake (high and variable EMG). EEG2 theta power was divided by EEG2 delta power and graphed × EMG signal in the 50–150 Hz band to elucidate clusters of REM, NREM, and wake epochs. The remaining transitional epochs were scored manually. The accuracy of cluster scoring was validated per animal manually.

For longevity: 12-month post-surgical mice were scored in a 30-min period during the light cycle (~3 p.m.) and utilized to bin power data by wake *vs*. sleep vigilance state.

For representative data (baseline vs. recovery): full 12-h dark cycle (7 p.m. –7 a.m.) and light cycle (7 a.m.–7 p.m.) data were scored for REM, NREM, and wake stages from both the baseline and recovery after SD recordings. The proportion of epochs per stage was calculated for the full 12-h recordings, as well as in 2-h bins. Arousals [1–2 wake epochs (10–20 s) during

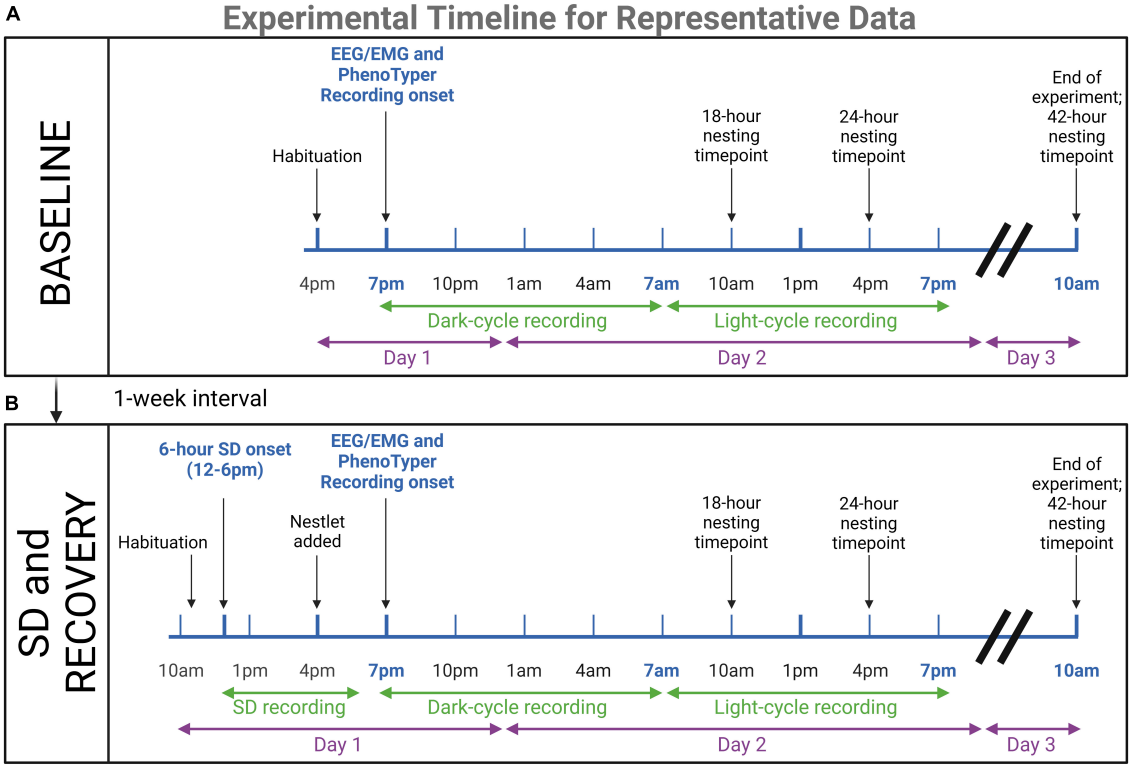


FIGURE 2
Timeline for representative data experiments. Young (n = 4) and aged (n = 4) mice underwent EEG/EMG recordings in PhenoTyper home cages to assess dark and light cycle sleep patterns and nesting. **(A)** 24-h baseline recordings were conducted, with 3 h of habituation prior to recording onset at 7 p.m. At 10 a.m. on the third day, mice were removed from the cage and potentiostats were unplugged. Nesting timepoints were 18, 24, and 42 h. **(B)** One week later, mice were tested again in the same paradigm with the addition of a 6-h sleep disruption (SD) prior to the recording. This SD was conducted in the home cages with aversive stimuli. Mice were plugged into potentiostats and habituated 1-h before the onset of SD (12–6 p.m.). The nestlet was added at 4 p.m. to keep nesting time points consistent. The immediate recovery period was recorded for 24-h to compare to baseline sleep patterns. The results of these experiments are presented as representative data of the utility of the concurrent electrophysiological and behavioral technique for age assessments (refer to **Figures 6**, **7**). Created with BioRender.com.

sleep bouts] and micro-sleep shifts [1–2 sleep epochs (10–20 s) during wake bouts] were quantified during light-cycle baseline and recovery recordings and normalized to the sum of sleep and wake bouts, respectively. A bout was defined as 3+ continuous epochs, and the occurrence of 3+ continuous epochs of the opposite stage indicated the end of the current bout. NREM and REM were grouped for the calculation of sleep bouts.

For SD: the 6-h recording was scored, the proportion of epochs per stage calculated, and REM, NREM, and wake data were compared to the equivalent timespan during the baseline recording.

3.8. Statistical analysis

GraphPad Prism 9 was utilized for the generation of graphs and statistical analyses. Data are presented as mean ± SEM and best-fit line as ±95% CIs. Repeated measures of ANOVA and paired *t*-tests were utilized to assess longevity, Barnes maze analyses, nesting, sleep disruption, and movement with potentiostats. Unpaired *t*-test was utilized for comparing arousals and stage shifts across age. Where necessary, multiple comparisons were controlled for with Tukey (*F*-test statistics) and with Holm–Šídák (*T*-test statistics) *post hoc* tests. Linear and non-linear regression (exponential one-phase decay, least squares fit) were utilized to fit lines for neuronal frequency × locomotion. Two mice were graphed, but were omitted from the 12-month post-surgery statistical analysis due to the inability to distinguish sleep vs. wake; explained further in results text. In the neuronal frequency X locomotion dark-cycle data, 16 data points (out of 576 total) were excluded due to abnormally high EEG amplitudes in one mouse near the onset of recording; this did not persist. Two mice used in initial optimization were omitted from locomotion- and Barnes maze-associated EMG analyses, due to EMG signal degradation; the EEG2 signal reported was not affected. All values reported in the text (i.e., effect sizes) are mean ± SEM. Exact statistical tests and the number of individuals are reported in the results section and figure legends.

4. Results

We tested the application of *in vivo* wireless EEG/EMG recording during behavioral testing in mice as a proof-of-concept for assessing simultaneous electrophysiological and behavioral changes critical in aging and neurodegenerative research. Herein, we demonstrate the viability of recordings longitudinally, as well as during a variety of behavioral tasks and interventions, including home-cage locomotor assessment, activities of daily living, cognitive testing in the Barnes maze, and sleep staging at baseline and after automated sleep disruption (Figure 1).

4.1. Advantage #1: Quality signal can be ensured for longitudinal recordings up to 9 months post-surgery, with the potential for 12-month post-surgical sleep staging

Twelve 3.5-month-old mice were implanted with EEG/EMG headcaps to assess the potential for longitudinal recordings with the system. Locomotor home cage and EEG/EMG recordings were collected from 8 mice at 1, 5, and 9 months post-surgery, and 7 mice at the 12-month post-surgery at approximately 4-5, 8, 12, and 16 months of age, respectively (Figure 3). Four mice died from health-related issues (one was excluded only from the 12-month post-surgical analyses), and 1 had an issue with headcap stability after 8-month post-surgery testing, which could have been circumvented with re-application of cement. Table 1 annotates these details, and section "4.5 Pitfalls, limitations, troubleshooting, and alternatives" provides tips to improve stability/longevity. During recordings, mice adjusted well to the added weight of the potentiostat plugged in (∼3.4 g), exhibiting no differences in mobility when compared to mice without the potentiostat (refer to Supplementary Figure 3A).

The EEG power by frequency was generated on the anterior electrode for alpha (8–13 Hz), beta (13–30 Hz), gamma (35–44 Hz), delta (0.5–4 Hz), and theta (5.5–8.5 Hz) waveforms and normalized to the full channel signal during a predominant wake-period (Figure 3A). No overall significant differences were detected by 1, 5, and 9 months post-surgical time: $F = 0.14$, DFn, DFd = 1.55, 10.87, $P = 0.82$. Frequencies significantly differed from each other ($F = 31.79$, DFn, DFd = 1.15, 8.06, P = 0.0004), with no frequency × post-surgery interaction effect [$F = 2.54$, DFn, DFd = 1.80, 12.62, $P = 0.12$ (two-way repeated measures ANOVA)]. The *post hoc* analysis determined a significantly lower percentage of alpha power at 9 vs. 1 month post-surgery time ($P = 0.014$); no other comparisons differed (controlled for multiple comparisons with Tukey *post hoc* test; Figure 3A). No sex differences were detected in EEG frequency distribution at each post-surgery assessment (1 month: $F = 0.045$, DFn, DFd = 1, 6, $P = 0.84$; 5 months: $F = 0.26$, DFn, DFd = 1, 6, $P = 0.63$; 9 months: $F = 0.023$, DFn, DFd = 1, 6, $P = 0.88$; two-way repeated measures ANOVA; Figure 3A).

The full power for EEG2 (anterior electrode; $F = 1.06$, DFn, DFd = 1.28, 8.94, $P = 0.35$; Figure 3B), EEG1 (posterior electrode; $F = 0.64$, DFn, DFd = 1.31, 9.19, $P = 0.49$; Figure 3C), and EMG ($F = 0.00038$, DFn, DFd = 1.74, 12.17, $P = 0.99$; Figure 3D) were assessed (one-way repeated measures ANOVA), also distributing no significant differences at 1, 5, and 9-months post-surgery. No sex differences were detected for full power measures either (EEG2: $F = 0.15$, DFn, DFd = 1, 6, $P = 0.71$; EEG1: $F = 0.28$, DFn, DFd = 1, 6, $P = 0.62$; and EMG: $F = 0.39$, DFn, DFd = 1, 6, $P = 0.56$; two-way repeated measures ANOVA; Figures 3B–D). These data demonstrate the preservation of headcap signal integrity up to 9 months

10.3389/fnagi.2022.952101

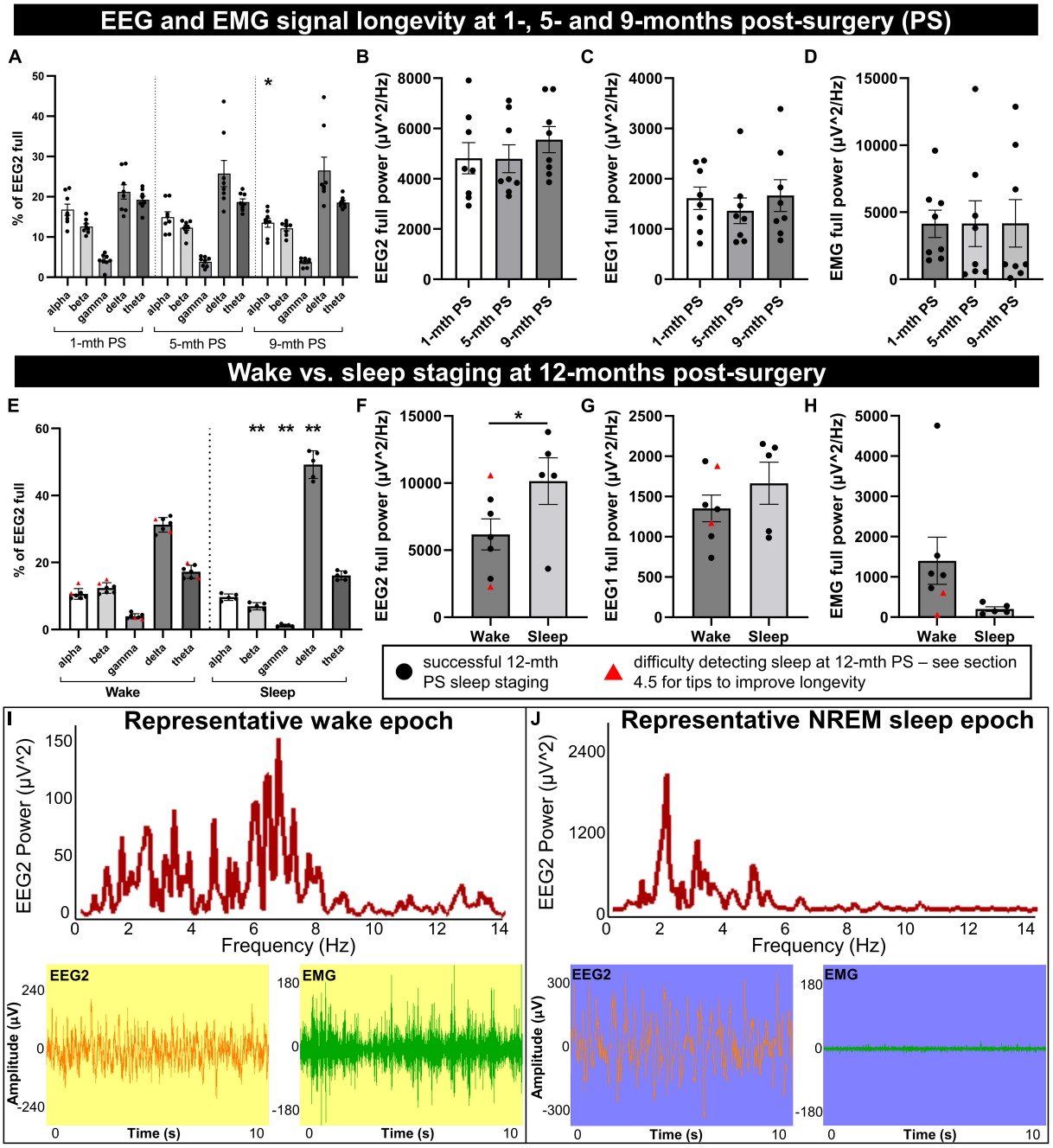


FIGURE 3

Longevity of EEG signal integrity until 9 to 12 months post-implantation. Mice were implanted with EEG/EMG headcaps, and recordings were collected at 1, 5, and 9 months post-surgery (PS) to look at signal longevity (n = 8), and sleep was staged at 12 months post-surgery (PS) (n = 7). Power was generated in alpha (8–13 Hz), beta (13–30 Hz), delta (0.5–4 Hz), theta (5.5–8.5 Hz), and gamma (35–44 Hz) bands for the anterior EEG electrode (EEG2), as well as full range (0.5–500 Hz) for EEG2, EEG1 (posterior) and EMG. **(A)** No overall effects of post-surgical time were detected on the proportion of neuronal frequency bands; differences between bands were significant on ANOVA. One significant *post hoc* comparison was detected between alpha banded power at 9 vs. 1-month post-surgery (P = 0.014). The full power of EEG2 **(B)**, EEG1 **(C)**, and EMG **(D)** channels did not significantly change between 1, 5, and 9 months post-surgery. At 12 months post-surgery, power was split by vigilance state for wake vs. sleep. **(E)** During sleep, there were proportionally fewer beta and gamma waves, and more delta than wake. Full EEG2 power was significantly increased **(F)**, EEG1 trended (P = 0.055) to increase **(G)**, and EMG trended (P = 0.097) to decrease **(H)**, in sleep compared to wake. Sleep was difficult to detect in two mice by cluster or manual staging (marked by red triangles), likely due to consistently low EMG. **(I,J)** Representative frequency distribution of wake and NREM sleep in 10 s epochs, with the corresponding EEG and EMG traces, demonstrate low amplitude/mixed frequency and high amplitude delta waves, respectively. **Supplementary Figure 4** documents raw traces for REM, NREM, and wake and epoch frequency distribution for REM. Data are mean ± SEM. Two-way **(A)** and one-way **(B–D)** repeated measures ANOVA, multiple comparisons controlled for with Tukey *post hoc* test; multiple paired t-test with multiple comparisons controlled for with Holm–Šídák method **(E)**, paired t-test **(F–H)**. *P < 0.05, **P < 0.01.

post-surgery. However, the loss of EMG signal is observed with increasing post-surgical time as the EMG leads can lose signal integrity due to the wear and tear of not being fixed in place, resting in the nuchal muscles. This does not necessarily preclude the differentiation of sleep and wake stages, which we demonstrate can be reliably measured at 12 months post-surgery in the same mice (**Figures 3E–I**).

Utilizing EEG and EMG signals, sleep vs. wake stages were identified in 10-s epochs for a 30-min period during the light-cycle, in 7 mice at 12 months post-surgery. Sleep was associated with significant reductions in the proportion of beta and gamma EEG2 signal (t = 7.43, df = 4, P = 0.0070 and t = 6.62, df = 4, P = 0.0081, respectively), increased delta (t = 9.58, df = 4, P = 0.0033), and no changes in alpha or theta (t = 0.92, df = 4, P = 0.65 and t = 0.84, df = 4, P = 0.65, respectively; **Figure 3E**) by multiple paired t-test (controlled for multiple comparisons with Holm–Šídák). Notably, increased delta frequency is characteristic of sleep (Brown et al., 2012). The full power of EEG2 significantly increased in sleep vs. wake (paired t-test: t = 4.30, df = 4, P = 0.013; **Figure 3F**), likely due to high amplitude delta waves, with a trend to increase in EEG1 full power (paired t-test: t = 2.68, df = 4, P = 0.055; **Figure 3G**). EMG full power trended to decrease in sleep vs. wake (paired t-test: t = 2.16, df = 4, P = 0.097; **Figure 3H**). Two of the mice, indicated by red triangles in **Figures 3E–H**, had no detectable sleep stages, which we concluded was due to the degradation of the EMG signal, and because the signal was consistently low (see **Figure 3H**), we were unable to differentiate sleep from wake and these mice were omitted from the paired statistical analyses in **Figures 3E–H**.

Using the actigraphy-like locomotion data collected in the PhenoTypers, we can differentiate sleep and wake stages by velocity ($\leq$0. 1 cm/s for sleep) with 86.39 $\pm$ 3.35% (n = 2) accuracy to EMG, demonstrating the utility of acquiring locomotor and EEG data, as well as an alternative staging method to ensure the longevity of recordings.

Representative frequency distribution of wake and NREM sleep epochs at 12 months post-surgery demonstrate mixed frequency low-amplitude EEG signal and high-amplitude, low-frequency EEG signal, respectively (**Figures 3I, J**). REM sleep (high theta:delta ratio and low EMG) is also detectable at 12 months post-surgery (refer to **Supplementary Figure 4A** for representative EEG and EMG traces for all three stages and **Supplementary Figure 4B** for REM epoch frequency distribution).

When comparing neuronal power between animals, it may be best to plan for consistent post-surgery times and with normalization of power with another readout (i.e., as a percent of total EEG power, to locomotion or another behavioral readout). However, this data demonstrates that up to 9 months post-surgery, the quality of EEG and EMG signals was not significantly degraded and was sufficient for sleep staging, and can be successfully extended to 12 months post-surgery with proper troubleshooting (see section "4.5 Pitfalls, limitations, troubleshooting, and alternatives").

For the EEG/EMG analyses in sections 4.2, 4.3, and 4.4, data from the four mice dropped from the longevity analyses was utilized along with 4 additional mice from our optimization and pilot studies (refer to **Table 1**). In light of no post-surgical or sex differences determined in signal longevity, female and male mice, and those with different post-surgical times were grouped together.

4.2. Advantage #2: Association of neuronal frequency with locomotor activity

We next binned neuronal power by frequency for delta (0.5–4 Hz), theta (5.5–8.5 Hz), alpha (8–13 Hz), and beta (13–30 Hz) waves, to assess for potential relationships with locomotion in PhenoTyper home-cage recordings in 8 mice; 2 mice were omitted from EMG analyses (refer to **Table 1**). Time-matched neuronal power (% of full power range) and velocity were graphed during sleep-dominant light-cycle and wake-dominant dark-cycle (**Figure 4**). Linear and non-linear regressions were conducted to fit power $\times$ locomotion relationships. An exponential one-phase relationship was observed with delta power and velocity during both light and dark cycles (df = 573, R^2 = 0.17, Y0 = 38.02, plateau = 28.83, K = 1.78 and df = 557, R^2 = 0.52, Y0 = 43.10, plateau = 22.56, K = 1.62, respectively; **Figure 4A**), demonstrating high delta power during periods of low velocity (likely sleep) and tapering off at higher velocities, especially during the dark-cycle. Delta waves are dominant during NREM and slow wave sleep (**Figure 4I**) (Brown et al., 2012).

No velocity $\times$ theta relationship was observed during the light cycle. We fit a one-phase line during the dark cycle (df = 557, R^2 = 0.11, Y0 = 18.20, plateau = 21.44, K = 1.85; **Figure 4B**); a slight increase in the dark cycle perhaps demonstrated the relevance of theta waves in REM as well as wake-associated memory processes (**Figure 4I**) (Brown et al., 2012). Alpha power did not alter with velocity in the light cycle; yet, interestingly, exhibits a positive linear relationship during the dark cycle (R^2 = 0.11, slope = 2.87, Y-intercept = 13.39, F = 501.4, DFn, DFd = 1, 558, P < 0.0001 for slope significantly non-zero; **Figure 4C**). This is likely a result of lower delta power and lesser sleep states in these higher velocity data points during the dark cycle, leading to a higher proportion of alpha, but not higher total alpha power. No linear relationship was observed between locomotion and total alpha power during the dark cycle (R^2 = 0.00012, slope = −4.81, Y-intercept = 1151, F = 0.067, DFn, DFd = 1, 558, P = 0.80). Alpha waves are increased during wakefulness and thought processes and, therefore, indicative of greater activity. They can also associate

10.3389/fnagi.2022.952101

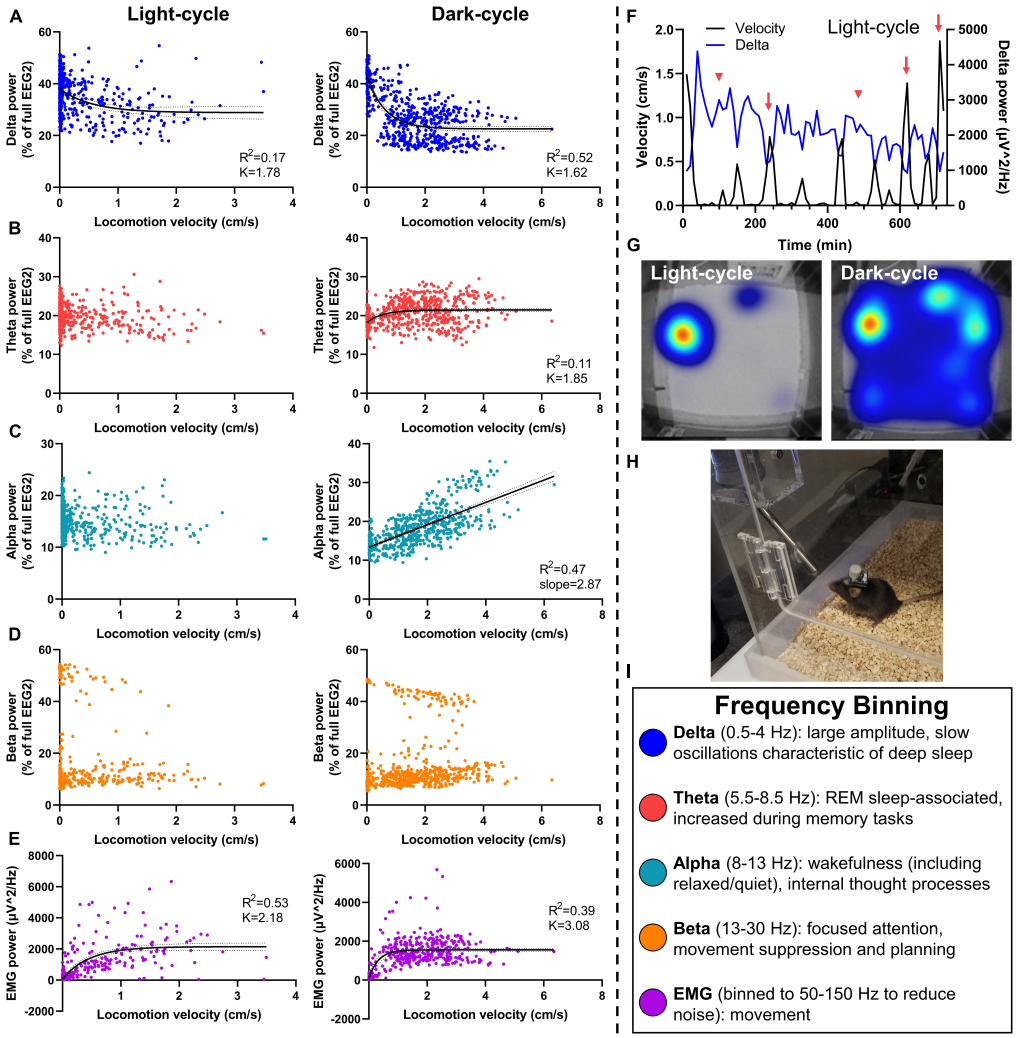


FIGURE 4

Relationship of neuronal frequency with locomotor activity. Mice were assessed in PhenoTypers with concurrent EEG/EMG recordings over 24 h in the light and dark cycles. Velocity was aligned with relative EEG frequency bands (normalized to full power) in 10-min bins to assess relationships between locomotion and sleep behavior with neuronal power. **(A)** Delta power was negatively associated with movement in light and dark cycles, demonstrating notably high delta during low-velocity recordings indicative of NREM and slow wave sleep, and a taper from delta power with high locomotion. **(B)** Theta power demonstrated a slight relationship in the dark cycle only. **(C)** A linear velocity × alpha relationship was detected in the dark-cycle only. **(D)** No velocity × beta relationships were observed. **(E)** EMG (50–150 Hz) exhibits a positive relationship with velocity in both light and dark cycles. **(F)** Light-cycle graph of velocity and delta power over time in a representative mouse demonstrates the increased delta power during sleep bouts (red arrowheads), and decreased delta power during movement (red arrows). **(G)** PhenoTyper heatmap of movement in the light and dark cycle. **(H)** Mice with EEG/EMG headcap and potentiostat exhibit normal exploratory and locomotor behavior in PhenoTyper (refer to **Supplementary Figure 3A**). **(I)** EEG frequency bins are listed with relationships to behavior. Data are 10-min bins of power and locomotor activity, with the line of best fit ±95% CIs, from n = 8 mice **(A–D)** or n = 6 mice **(E)**. Linear and non-linear regression (exponential one-phase decay, least squares fit).

with quiet wakefulness in which there is rest and reduced movement but not sleep (**Figure 4I**) (Brown et al., 2012), perhaps contributing to differing alpha × velocity correlations in the light- and dark cycles.

Beta power had no association during the light- or dark cycle (**Figure 4D**), indicative of the bimodal function of beta for movement inhibition, as would be observed with sleep onset, compared with attention and planning during wakefulness (**Figure 4I**) (Brown et al., 2012; Kropotov, 2016). Finally, we fit a one-phase decay line of EMG with velocity during the light and dark cycles (df = 429, R^2 = 0.53, Y0 = 73.88, plateau = 2,149, K = 2.18 and df = 413, R^2 = 0.39, Y0 = 82.20, plateau = 1554, K = 3.08, respectively; **Figure 4E**).

A representative time-based graph of velocity and delta power in the light cycle demonstrates increased delta during sleep bouts (low velocity) and vice versa (**Figure 4F**). When undisturbed, mice spend most of the light cycle in their nest (red–yellow hot-spot in the light cycle heat map), compared

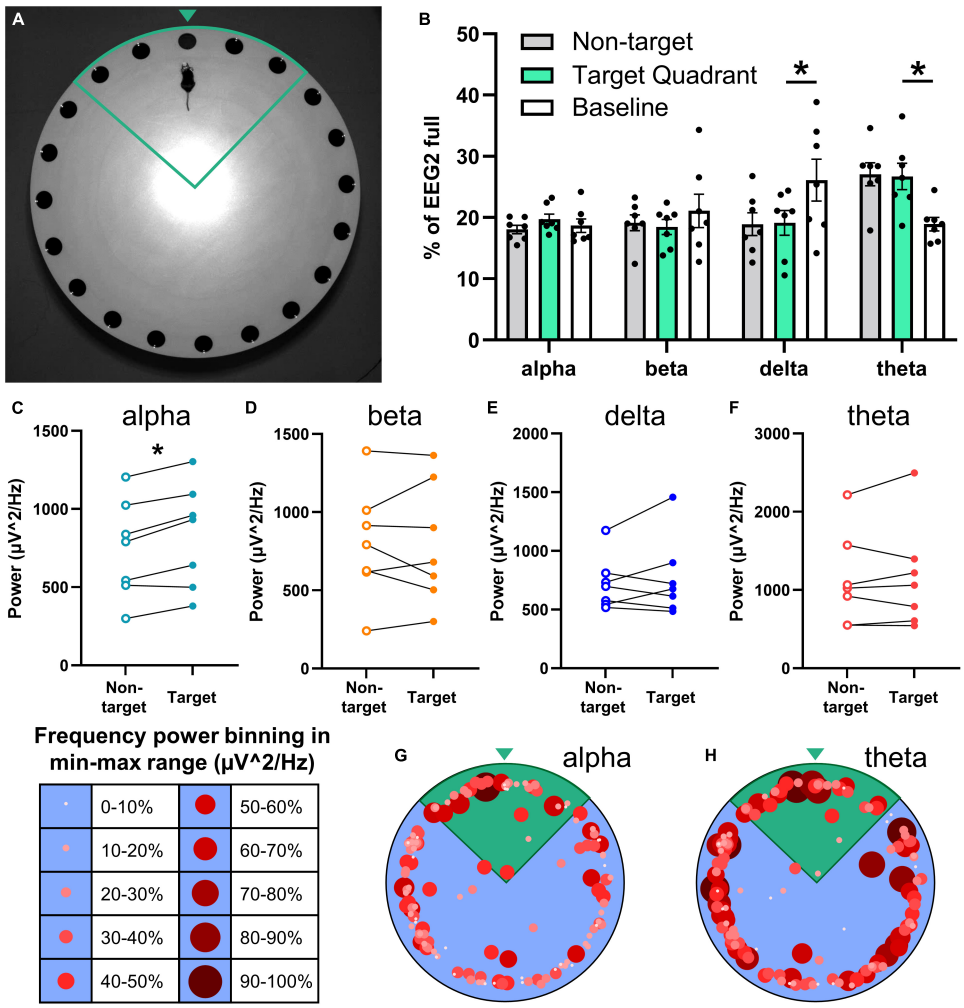


FIGURE 5
Relationship of the neuronal frequency with Barnes maze performance. EEG/EMG mice (n = 7) were tested in the Barnes maze as a proof-of-concept of combining recordings with cognitive testing. Mice learned the task successfully (refer to **Supplementary Figures 5A,B**) across four trial days (two trials per day) and were tested for spatial memory 2 days later in a probe trial in which the previously acquired escape location was blocked. **(A)** Mice with the potentiostat plugged-in navigate the Barnes maze successfully (refer to **Supplementary Figure 3B**), demonstrated here in the target quadrant (green outline) approaching the target zone (green arrowhead). **(B)** Neuronal power was generated in 2-s intervals, normalized to full EEG signal, and binned by mouse location in the Barnes maze (non-target and target quadrant), as well as at baseline (not performing task). While searching for the escape in the target quadrant, mice exhibit significantly lower delta power and significantly greater theta power than at baseline. Paired power comparisons of non-target vs. target quadrant determined significantly greater alpha power **(C)** in the target zone, and no differences for beta **(D)**, delta **(E)**, theta **(F)**, and EMG (refer to **Supplementary Figure 5C**) during the task. **(G,H)** Alpha and theta power were binned as a proportion (10% intervals) of the total power range recorded at each 2-s interval and graphed by size, color, and xy location in the Barnes maze. These heatmaps demonstrate higher alpha power in the target quadrant (green), and high theta throughout the maze. Representative heatmaps were generated in the same mice; positional data is the same for both. Data are mean ± SEM **(B)** or paired before-after comparisons **(C–F)**. Multiple paired t-tests with multiple comparisons were controlled for with the Holm–Šídák method. *$P < 0.05$.

to high activity during the dark-cycle (broader heat map distribution in blue) (**Figure 4G**). PhenoTypers facilitate a controlled environment for consistent locomotor and EEG/EMG recordings, in which the mice are comfortable and have *ad libitum* access to food and water (**Figure 4H**). Preferably, these recordings are conducted in a dedicated testing room as the additional noise, vibrations, and electrical/Bluetooth interference from traffic in normal housing rooms can interfere with EEG/EMG signals.

4.3. Advantage #3: Association of neuronal activity with cognitive performance in the Barnes maze

To probe the viability of the wireless EEG/EMG system with additional behavioral assays, we assessed headcap mice in the Barnes maze, a test of spatial learning and memory (**Figure 5**). Mice successfully learned the location of the

escape box based on spatial cues around the room (refer to **Supplementary Figures 5A,B**). Potentiostats were not plugged-in during learning trials in the present study because we anticipated the escape hole was not large enough (5 cm diameter) to accommodate the potentiostat. Utilizing larger escape holes or additional training time to help guide mice into the hole may facilitate EEG/EMG data collection during all trials. Three days following the last learning trial, potentiostats were plugged into mice headcaps, and mice were tested in a spatial memory probe in which the escape box was blocked. In this trial, mice were assessed for the percentage of time spent searching in the target quadrant during the 5-min trial (35.78 ± 4.10%; n = 7). Potentiostats did not affect mobility in the Barnes maze (refer to **Supplementary Figure 3B**). Representative image of a mouse with wireless EEG/EMG unit in the Barnes maze headed toward the escape location with target quadrant (green outline) and zone (green arrowhead) indicated in **Figure 5A**.

Neuronal power was analyzed by frequency in 2-s bins and aligned with the location in the Barnes maze in 7 mice. Frequency distribution was generated for neuronal power when mice were in the Barnes maze searching in non-target zones, in the target quadrant, and baseline activity before and after the task (**Figure 5B**). Target quadrant vs. baseline comparisons determined significantly less slow wave sleep-associated delta waves (t = 3.32, df = 6, P = 0.047) and greater theta power (t = 3.84, df = 6, P = 0.034) when mice were searching for the escape; alpha (t = 0.81, df = 6, P = 0.70) and beta (t = 0.76, df = 6, P = 0.70) waveforms did not change (multiple paired t-test; controlled for multiple comparisons with Holm–Šídák).

We assessed total power by frequency to further probe the potential relationship of neuronal power with performance in the Barnes maze. We observed a significant increase in alpha power in the target compared to non-target zones (t = 4.59, df = 6, P = 0.015; **Figure 5C**), but not in beta (t = 0.086, df = 6, P = 0.93; **Figure 5D**), delta (t = 0.80, df = 6, P = 0.84; **Figure 5E**) or theta (t = 0.50, df = 6, P = 0.87; **Figure 5F**) waves (multiple paired t-test; controlled for multiple comparisons with Holm–Šídák). EMG power (in 50–150 Hz) did not change in the target quadrant vs. non-target (n = 5; t = 1.36, df = 4, P = 0.25; paired t-test; refer to **Supplementary Figure 5C**).

In sum, higher alpha power was observed when mice were in the target quadrant and close to checking the correct hole, and higher theta power was observed throughout the task; these frequencies are associated with memory, thought, and attention (**Figure 4I**) (Brown et al., 2012; Kropotov, 2016). To visualize the association of neuronal frequency with the location in the Barnes maze, the power range of alpha and theta waves were separated into 10% bins and graphed by the xy location (**Figures 5G, H**). These representative images demonstrate higher selectivity of increased alpha neuronal power closer to the escape location, whereas theta power was less discriminate but high throughout the task. These data demonstrate the feasibility of EEG/EMG recordings simultaneous to behavioral testing in rodents. We anticipate that the combination of these techniques can help pinpoint electrophysiological alterations as they relate to cognitive decline and behavioral alterations across age and during the course of neurodegenerative disease progression.

4.4. Representative data: Sleep staging in young vs. aged mice at baseline and following sleep disruption

We assessed young (4–5 months) and aged (9–10 months) mice with EEG/EMG headcaps to measure neuronal activity and locomotion in Noldus PhenoTypers across a full circadian cycle (24 h) at a baseline starting from the onset of their dark cycle (7 p.m.). One week later, the same mice were subjected to a 6-h sleep disruption (SD; 12–6 p.m.) *via* a loud tone and bright light inside each PhenoTyper cage, set on a procedurally randomized timer. Mice were measured for the next 24-h cycle after SD to assess the recovery period (refer to **Figure 2** for timeline). Sleep was staged across the entire recording at baseline and in the recovery period as a representative of alterations across age and after SD (**Figure 6**).

At baseline during the dark cycle, young and aged mice spend the majority of time awake and start to sleep more as the onset of the light cycle approaches (7 a.m.; **Figure 6A**). During the baseline light cycle, mice sleep approximately 60% of the time, and young mice trend to more REM and less NREM sleep than aged mice (**Figures 6B, C**). Immediately following SD, mice were tested again to assess their sleep activity during the recovery period. More variability between sleep and wake states was observed during the recovery dark cycle, and young mice trended to sleep more compared with the baseline (**Figure 6D**). Slight decreases in sleep time were observed in the recovery period for young and aged mice, with less REM sleep in young mice (**Figures 6E, F**). **Figures 6G, H** summarizes these trends.

We analyzed for wake arousals detected during sleep-dominant periods in the light cycle at baseline and in the recovery period. Arousals did not significantly differ between young and aged mice at baseline (t = 0.86, df = 6, P = 0.60) or recovery (t = 0.97, df = 6, P = 0.60; multiple unpaired t-tests with multiple comparisons controlled for with Holm–Šídák method; **Figure 6I**), though it may increase in aged mice after SD (t = 1.84, df = 3, P = 0.16, paired t-test; effect size: 2.05× ± 0.59 increase). We then analyzed for micro-sleep episodes detected during wake-dominant periods in the light cycle, which were determined to be significantly more in aged mice at baseline (t = 3.63, df = 6, P = 0.022), but there were no differences in the recovery period (t = 0.34, df = 6, P = 0.75; **Figure 6J**).

10.3389/fnagi.2022.952101

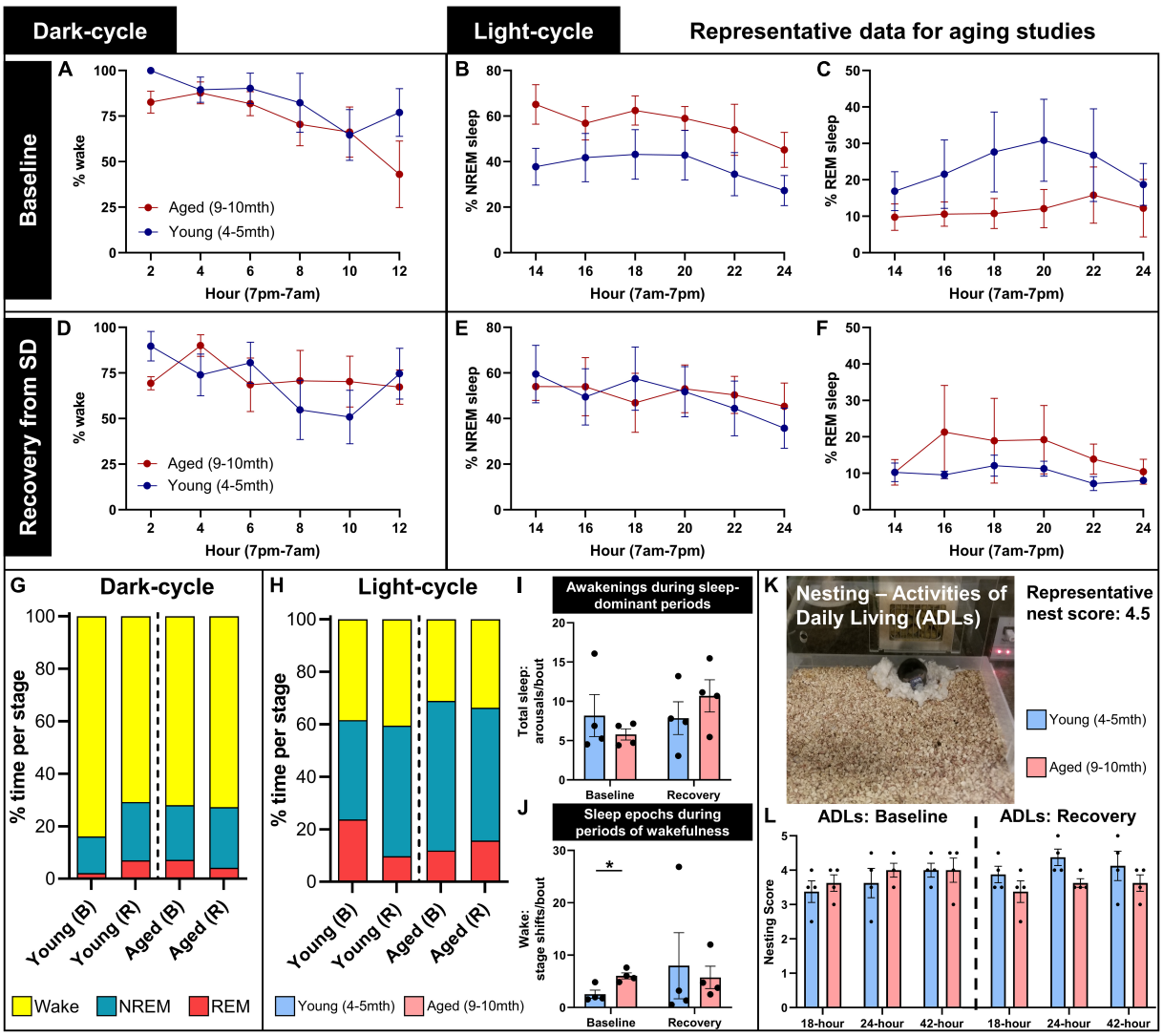


FIGURE 6

Representative data for aging studies on sleep and home-cage assessments. Young (4–5 months; $n = 4$) and aged (9–10 months; $n = 4$) mice were recorded for a full circadian cycle in PhenoTyper home cages with EEG/EMG recording from wireless headcaps. Recordings were conducted at a baseline, and in a recovery period immediately following 6-h sleep disruption (SD), 1 week later (refer to Figure 2 for timeline). Sleep was staged as wake, NREM, and REM in 10-s epochs from EEG/EMG recordings, split into 12-h dark and light cycles and analyzed in 2-h intervals. **(A)** During the baseline dark cycle, young and aged mice spend a majority of their time awake, and start to sleep more near the onset of the light cycle. **(B,C)** During the baseline light cycle, young mice trend to less NREM and more REM sleep than aged mice. **(D–F)** In the recording period immediately following SD (indicated here as recovery), all mice trend to more sleep during dark-cycle, and young mice trend to more NREM and less REM sleep during light cycle, compared to baseline. **(G,H)** Summary of the distribution of wake (yellow), NREM (blue), and REM (red) in young and aged mice in baseline (B) and recovery (R) recordings, for dark- and light-cycles. **(I)** During light-cycle, no significant differences in sleep arousals were detected between young and aged mice at baseline and recovery; **(J)** however, aged mice exhibited significantly more sleep epochs during periods of wakefulness at baseline ($P = 0.022$), indicating increased sleep fragmentation with age. **(K,L)** Nest building was scored at 18, 24, and 42 h as a measure of activities of daily living during baseline and in the post-SD recovery period, demonstrating the trend to reduced nest complexity in aged mice in the recovery period, compared with young mice ($P = 0.081$). Data are mean ± SEM **(A–F,I,J,L)** and mean **(G,H)**. Multiple unpaired *t*-tests with multiple comparisons controlled with the Holm–Šídák method **(I,J)**; two-way repeated measures ANOVA **(L)**. *$P < 0.05$.

These data demonstrate increased sleep fragmentation with age.

Nesting was assessed during circadian recordings as a measure of activities of daily living. A representative image demonstrates a mouse with a headcap and potentiostat in the PhenoTyper home cage with a complex nest built (score: 4.5/5; Figure 6K). Nest complexity was scored at 18, 24, and 42 h during baseline and post-sleep disruption (recovery) recordings in young and aged mice. At baseline, young and aged mice trend to higher nest scores over time ($F = 4.59$, DFn, DFd = 1.27, 7.62,

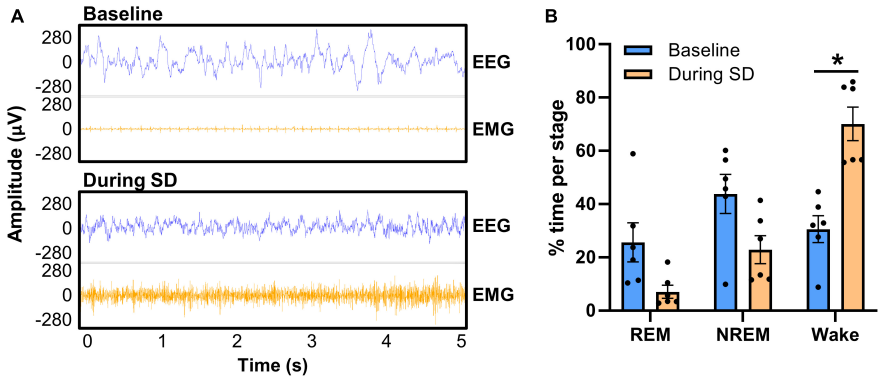


FIGURE 7

Confirmation of PhenoTyper-mediated sleep disruption in EEG. Mice (n = 6) underwent a baseline recording and an acute, 6-h sleep disruption (SD; 12–6 p.m.) 1 week later (refer to **Figure 2** for timeline). SD involved an intermittent tone and light from the top-unit of the PhenoTyper so that each mouse was exposed to the same stimuli. The interval between both tone and light was procedurally randomized to prevent habituation. **(A)** Neuronal activity was recorded by EEG/EMG headcaps during SD and resembles wakefulness, with higher frequency EEG and EMG than sleep (baseline). Heartbeat can be detected in EMG when the mice are at rest [as in the baseline EMG in panel **(A)**], but is not always present. **(B)** EEG/EMG data was staged as NREM, REM, and wake in the same time frame at baseline and SD, demonstrating a reduction in the time spent asleep and a significant increase in wakefulness (P = 0.022). Data are mean ± SEM. Multiple paired t-tests with multiple comparisons were controlled for with the Holm–Šídák method. *$P < 0.05$.

P = 0.060), with no differences between ages (F = 0.30, DFn, DFd = 1, 6, P = 0.60). During the recovery period, aged mice trend to lower nest complexity than young mice (F = 4.39, DFn, DFd = 1, 6, P = 0.081; two-way repeated measures ANOVA; **Figure 6L**). We anticipate that this is indicative of a loss of activities of daily living after lifestyle stressors, and how over-age these stressors and a diminished ability to compensate for them can potentially exacerbate behavioral deficits (Webb et al., 2018).

We next confirmed in EEG that sleep was disrupted utilizing the tone and bright light in the PhenoTypers (**Figure 7**). Representative 5-s EEG and EMG traces demonstrate high amplitude slow waves during sleep (baseline) with tonic and low EMG amplitude, in contrast to the higher frequency, lower amplitude neuronal activity, and high EMG signal during SD, indicative of wakefulness (**Figure 7A**). Six out of the eight mice in **Figure 6** were utilized for SD confirmation; the other 2 mice had battery issue causing data loss for half of the SD period, though they still exhibited disrupted sleep (refer to section "3.3 EEG/EMG recording and alignment with behavioral assessments," step 3 for tips on ensuring battery quality). Assessment of REM, NREM, and wake staging during the 6-h SD, compared with time-matched baseline (multiple paired t-test with multiple comparisons controlled for with Holm–Šídák method), indicates significantly more wakefulness during the PhenoTyper-mediated SD (t = 4.34, df = 5, p = 0.022), with 55.29 ± 10.10% (n = 6) less time spent asleep (**Figure 7B**). Time in the REM (t = 2.10, df = 5, p = 0.17) and NREM (t = 1.79, df = 5, p = 0.17) stages trended to less during SD compared with baseline (**Figure 7B**).

These data indicate that randomized auditory and light stimuli in PhenoTyper cages are sufficient to disrupt sleep. We propose this method of sleep disruption recapitulates fragmenting of sleep that occurs with age (Scullin and Bliwise, 2015), instead of complete deprivation, and that an impaired ability to recover after sleep loss is a critical readout to understand how mounting sleep deficits may impact neurodegenerative diseases (Musiek and Holtzman, 2016).

4.5. Pitfalls, limitations, troubleshooting, and alternatives

1. The headcap procedure is feasible but is more challenging in female mice and mice under 3 months due to size. Consequently, the placement of the headcap needs greater precision to ensure that the anterior and posterior electrode screws are above the cortex.
2. We have observed 2 mice (out of 58 successful implantations; 3.45%) lose a headcap, at 5 and 9 months post-surgery. Both were identified as having reduced dental cement securing the headcap compared with mice that had surgeries at the same time. Dental cement can degrade over time, so it is critical to ensure coverage of screws and the base of the headcap at the time of surgery. A potential re-application of dental cement (every 3–6 months) applied under anesthesia may be necessary for longitudinal experiments in mice to prevent the headcap from getting displaced, or if EMG wires become exposed (re-cover connection to headcap to prevent further damage). This is quick (5–10 min/mouse), yet will depend on the time for the cement to set. Silicone coating of the exposed screwheads before dental cement

re-application may also preserve signal longevity, although this was not tested herein.

3. For longitudinal experiments, it is important to consider attrition of animals *via* death from non-headcap-related causes, when considering sample size, in addition to the potential loss or displacement of the headcap.
4. To help preserve EEG signal quality over time (refer to **Figure 3**), an alternative method is to utilize screws with a wire attached (8403, Pinnacle Technology Inc.), allowing soldering of the electrode to the headcap.
5. If there are issues in restraining the mice by the headcap for plugging-in in the potentiostat (i.e., trouble gripping the headcap from dental cement, agitated mice), a fast-acting and clearing inhalational anesthetic can be used. The investigator can utilize isoflurane restraint (5%) for 1–2 min, just for induction. Mice recover quickly but should be given ∼1-h post-anesthesia prior to data collection.
6. Utilizing the cluster scoring method in Sirenia® Sleep Pro leaves a few unscored transition epochs (refer to **Supplementary Figure 1**). Therefore, some manual scoring is necessary, but is quick.
7. The power data in the present study were generated utilizing the default frequency bins in Sirenia® Sleep Pro, which were well-aligned to behavioral and vigilance states, and advantageous for separating NREM and REM sleep epochs. However, an overlap exists between 8 and 8.5 Hz in our theta and alpha frequency bins, and there is potential for biologically relevant neuronal activity between 4 and 5.5 Hz being missed in the present analyses. Frequency binning can be edited in the software, allowing additional bands such as sigma [prominent in sleep spindles; 12–16 Hz (Holz et al., 2012)] and gamma (included by default; 35–44 Hz); albeit, the power of this low gamma range was low in our longevity analysis during wake and sleep states (refer to **Figures 3A, E**), and was hence omitted for subsequent analyses. Sirenia® Sleep Pro also allows manually defined sleep staging, which, for example, can allow the identification of slow wave sleep (by breaking down NREM by delta power) and quiet wakefulness (low EMG, low delta, and high alpha). This can be accomplished with manual, cluster, or threshold scoring, which allows users to define rules for scoring based on neuronal frequency power and EMG.
8. Alternatively, Freyburger et al. (2016) utilized shorter epochs (i.e., 4 s compared to 10 s in the present study) and analyzed power spectra by vigilance/sleep stages (REM, NREM, and wake) (McShane et al., 2010), which is advantageous in detecting short arousals and neuronal frequency changes by stage, respectively.
9. The exported data can be utilized to align other electrophysiological readouts to the behavioral phenotype, including coherence between the two EEG channels, phase synchronization, and cross-frequency coupling (Tort et al., 2009; Canolty and Knight, 2010; Hata et al., 2016). An alternative is to export EEG/EMG data acquired in Sirenia® as a .edf file for analysis in other software.
10. Alternate headcap systems: (1) Pinnacle offers the option for depth electrodes, advantageous in adding regional field potential recordings (i.e., from the hippocampus). (2) Spike Gadgets offers headcaps with a 32-channel capacity. (3) For other EEG systems, refer to Pinnell et al. (2015, 2016), Zayachkivsky et al. (2015), Vogler et al. (2017), Kent et al. (2018), Crouch et al. (2019), Medlej et al. (2019), Buenrostro-Jáuregui et al. (2020). A recent study conducted by Crouch et al. (2019) demonstrated the utility of combined EEG and behavioral assessments in phenotyping mouse models of AD. This technique provides greater temporal resolution than what we demonstrate herein; however, the possibility for longitudinal-based recordings as we describe is advantageous.
11. Chronic PhenoTyper-mediated sleep disruption was not tested with EEG/EMG recording. It is possible that mice adjust to the randomized tone and light stressors after continued exposure, in which case shorter intervals and longer tone/light lengths could reduce habituation, or alternatively, the usage of different sleep disruption techniques might be preferable (i.e., rotating bar and gentle handling) (Colavito et al., 2013).
12. The synchronization of behavioral and electrophysiological data herein was conducted utilizing the computer clock, with both software running on the same computer, and was accurate to the second. Higher precision can be attained with alternative methods: transistor–transistor logic (TTL) signals for data alignment are possible with both systems.
13. Though we observe no significant technical differences up to 9 months post-surgery, it is recommended to consider consistent post-surgical time in the study design.

5. Discussion

Herein, we extensively describe and vet a method for concurrent wireless EEG/EMG recording with untethered, freely moving testing, including alignment of neuronal oscillatory signals to the behavioral phenotype. A one-time headcap surgery provides sustained signal longevity of at least 9 months, with a high potential for 12-month post-surgery that would be useful for experimental modeling of sleep staging, especially for facilitating longitudinal experimentation, critical in aging and neurodegenerative disease research. We found that the headcap and the wireless potentiostat recording unit did not impede mouse locomotion, allowing synchronization with a variety of behavioral assays, which we demonstrate for locomotion and home-cage

PhenoTyper assessment, as well as cognitive performance in the Barnes maze.

We anticipate this procedure could be used in a majority of rodent behavioral tasks without the need to adapt the task to accommodate a tethered system, particularly advantageous in large arena assays (Barnes maze, open field), and without the concern of wire entanglement in walled assays (elevated plus maze, shelter/cognitive choice hole entries in PhenoTypers). However, some minor physical constraints exist with the wireless system, such as the size of the potentiostat impeding entry into the hole. This can be accommodated by using larger entry holes or longer training periods. Caution should also be considered in water-based assays, such as the Morris water maze as the device may not be waterproof (Pinnell et al., 2016).

For home-cage assessments, this system is particularly advantageous in that a full circadian rhythm could be captured in a single recording (with upward of 72 h before changing batteries), allowing the possibility for the recording of multiple days of sleep staging in one subject without intervention. This, paired with the consistent PhenoTyper-mediated sleep disruption paradigm we describe, can help elucidate sleep resilience in response to stress and deprivation, as well as how this is altered with disease and age (Brieva et al., 2021; Cai et al., 2021; Casale et al., 2021).

We identified suggestive relationships between neuronal frequency bins with locomotion and cognitive performance, which warrant further investigation. Delta power was associated with sleep bouts as evidenced by negative non-linear relationships with locomotion in light and dark cycles, and the proportion of alpha power was more dominant compared to other frequencies at higher locomotion states in the dark cycle, likely due to the reduced power of sleep-associated delta waves (Brown et al., 2012). While conducting the Barnes maze cognitive task, mice exhibited decreased delta and increased theta power compared to the baseline activity, indicative of a higher activity/vigilance state (Brown et al., 2012), and of memory and navigation (Buzsáki and Draguhn, 2004; Buzsáki and Moser, 2013; Jung and Carlén, 2021), respectively. Hippocampal and entorhinal cortical theta is typically associated with spatial memory performance, though the contribution of widespread cortical theta oscillations, including the prefrontal cortex (under our anterior electrodes), contribute to this cognitive processing (Buzsáki and Draguhn, 2004; Buzsáki and Moser, 2013; Jung and Carlén, 2021). Our results also suggest higher alpha power occurs near the target zone in the Barnes maze, which may relate to internally directed thought processes (Brown et al., 2012), as the mice search more precisely near where it anticipates the escape box. The contributions of EEG and neuronal waveforms to behavior have been previously reviewed (Brown et al., 2012; Kropotov, 2016; Helfrich and Knight, 2019; Beppi et al., 2021), whereas our results indicate the utility of acquiring oscillatory signals during behavioral testing in rodents. Notably, this technique facilitates the identification of electrophysiological alterations as they directly relate to the cognitive/behavioral phenotype in aging and neurodegenerative preclinical models, which may elucidate subtle or early markers of disease. The robustness of these observations could be important for future research, especially for predicting disease onset or progression.

The use of EEG as a physiological biomarker for clinical diagnosis and prognosis has become increasingly popular in recent years (Keizer, 2021). Significant differences in EEG activity have been described in neurodegenerative conditions such as AD, Parkinson's, and frontotemporal lobe dementia (Babiloni et al., 2011, 2020; Garn et al., 2015; Goossens et al., 2017). Specific EEG markers have been shown to correlate with AD severity and provide differential dementia diagnosis (Garn et al., 2015; Goossens et al., 2017). The value of EEG biomarkers extends to mood disorders such as depression (Kaiser et al., 2018; Dev et al., 2022), anxiety disorders (Pavlenko et al., 2009; Al-Ezzi et al., 2021), and neurodevelopmental disorders such as Attention-Deficit/Hyperactivity Disorder and Autism Spectrum Disorder (Wang et al., 2013; Matlis et al., 2015; Angelidis et al., 2016). The use of EEG in brain disorders may complement traditional diagnostic methods of structured interviews and questionnaires (Snyder et al., 2015; Keizer, 2021), and the ability to detect and characterize these biomarkers may prove essential for preclinical research.

Finally, EEG is frequently employed as a tool to understand human cognition (Helfrich and Knight, 2019; Beppi et al., 2021), indicating the importance of preclinical correlates to understand cognitive and electrophysiological impairments in both healthy aging and disordered brain states. We propose that the full cognitive assessments and freely moving electrophysiological data documented herein will improve the validity of research into a majority of brain disorders. A description of limitations and alternative approaches is included in section "4.5 Pitfalls, limitations, troubleshooting, and alternatives."

We anticipate the method described herein for concurrent EEG and behavioral testing in mice will be highly advantageous in preclinical research in a multitude of aging, neurological, and psychiatric disorders, including Alzheimer's and Parkinson's diseases, frontotemporal dementia, depression, and schizophrenia, all of which involve neuronal dysfunction and sleep dysregulation (Liu et al., 2004; Ju et al., 2014; Chahine et al., 2017; Fang et al., 2019; Waite et al., 2020; Wainberg et al., 2021), and will benefit from the alignment of neuronal oscillatory-to-behavioral deficits. Finally, the advent of investigating longitudinal readouts can help in the

identification of predictive EEG biomarkers of cognitive decline and the pathophysiology of diseases like Alzheimer's.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

All animal experiments were conducted in accordance with the ethical standards of the Canadian Council on Animal Care guidelines and approved by the Animal Care Committee of CAMH (Protocol #850).

Author contributions

CM and WY designed the study. CM and AT collected the data. CM, AT, and SG conducted data analysis. CM, EC, and WY wrote the manuscript. All authors edited and revised the manuscript, contributed to the article, and approved the submitted version.

Funding

Acknowledgment was made to the donors of Alzheimer's Disease Research, a program of BrightFocus Foundation, for support of this research (A2022016F; CM). This work was also supported by CAMH Discovery Fund (WY and CM).

Acknowledgments

The authors thank Dr. Abid Hussaini and Dr. Sean Hill for consultation on electrophysiological readouts, Dr. Thomas Prevot for assistance with PhenoTyper optimization, and Kelvin Tse for technical support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.952101/full#supplementary-material

References

Al-Ezzi, A., Kamel, N., Faye, I., and Gunaseli, E. (2021). Analysis of default mode network in social anxiety disorder: EEG resting-state effective connectivity study. *Sensors* 21:4098. doi: 10.3390/s21124098

Angelidis, A., van der Does, W., Schakel, L., and Putman, P. (2016). Frontal EEG theta/beta ratio as an electrophysiological marker for attentional control and its test-retest reliability. *Biol. Psychol.* 121, 49–52. doi: 10.1016/j.biopsycho.2016.09.008

Babiloni, C., Blinowska, K., Bonanni, L., Cichocki, A., De Haan, W., Del Percio, C., et al. (2020). What electrophysiology tells us about Alzheimer's disease: a window into the synchronization and connectivity of brain neurons. *Neurobiol. Aging* 85, 58–73. doi: 10.1016/j.neurobiolaging.2019.09.008

Babiloni, C., De Pandis, M. F., Vecchio, F., Buffo, P., Sorpresi, F., Frisoni, G. B., et al. (2011). Cortical sources of resting state electroencephalographic rhythms in Parkinson's disease related dementia and Alzheimer's disease. *Clin. Neurophysiol.* 122, 2355–2364. doi: 10.1016/j.clinph.2011.03.029

Beppi, C., Ribeiro Violante, I., Scott, G., and Sandrone, S. (2021). EEG, MEG and neuromodulatory approaches to explore cognition: current status and future directions. *Brain Cogn.* 148:105677. doi: 10.1016/j.bandc.2020.105677

Brieva, T. E., Casale, C. E., Yamazaki, E. M., Antler, C. A., and Goel, N. (2021). Cognitive throughput and working memory raw scores consistently differentiate resilient and vulnerable groups to sleep loss. *Sleep* 44:zsab197. doi: 10.1093/sleep/zsab197

Brown, R. E., Basheer, R., McKenna, J. T., Strecker, R. E., and McCarley, R. W. (2012). Control of sleep and wakefulness. *Physiol. Rev.* 92, 1087–1187. doi: 10.1152/physrev.00032.2011

Bubu, O. M., Brannick, M., Mortimer, J., Umasabor-Bubu, O., Sebastião, Y. V., Wen, Y., et al. (2017). Sleep, cognitive impairment, and Alzheimer's disease: a systematic review and meta-analysis. *Sleep* 40, 1–18. doi: 10.1093/sleep/zsw032

Buenrostro-Já,uregui, M., Rodríguez-Serrano, L. M., Chávez-Hernández, M. E., Tapia-de-Jesús, A., Mata-Luevanos, J., Mata, F., et al. (2020). Simultaneous monitoring of wireless electrophysiology and memory behavioral test as a tool to study hippocampal neurogenesis. *J. Vis. Exp.* e61494. doi: 10.3791/61494

Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745

Buzsáki, G., and Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat. Neurosci.* 16, 130–138. doi: 10.1038/nn.3304

Cai, Y., Wang, J., and Hou, L. (2021). Resilience improves the sleep quality in disabled elders: the role of perceived stress. *Front. Psychol.* 12:585816. doi: 10.3389/fpsyg.2021.585816

Canolty, R. T., and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends. Cogn. Sci.* 14, 506–515. doi: 10.1016/j.tics.2010.09.001

Carter, M. E., Yizhar, O., Chikahisa, S., Nguyen, H., Adamantidis, A., Nishino, S., et al. (2010). Tuning arousal with optogenetic modulation of locus coeruleus neurons. *Nat. Neurosci.* 13:1526. doi: 10.1038/nn.2682

Casale, C. E., Yamazaki, E. M., Brieva, T. E., Antler, C. A., and Goel, N. (2021). Raw scores on subjective sleepiness, fatigue, and vigor metrics consistently define resilience and vulnerability to sleep loss. *Sleep* 45:zsab228. doi: 10.1093/sleep/zsab228

Chahine, L., Amara, A., and Videnovic, A. (2017). A systematic review of the literature on disorders of sleep and wakefulness in Parkinson's disease from 2005-2015. *Sleep Med. Rev.* 35, 33–50. doi: 10.1016/j.smrv.2016.08.001

Colavito, V., Fabene, P. F., Grassi-Zucconi, G., Pifferi, F., Lamberty, Y., Bentivoglio, M., et al. (2013). Experimental sleep deprivation as a tool to test memory deficits in rodents. *Front. Syst. Neurosci.* 7:106. doi: 10.3389/fnsys.2013.00106

Crouch, B., Yeap, J. M., Pais, B., Riedel, G., and Platt, B. (2019). Of mice and motion: behavioural-EEG phenotyping of Alzheimer's disease mouse models. *J. Neurosci. Methods* 319, 89–98. doi: 10.1016/j.jneumeth.2018.06.028

Deacon, R. M. J. (2006). Assessing nest building in mice. *Nat. Protoc.* 1, 1117–1119. doi: 10.1038/nprot.2006.170

Dev, A., Roy, N., Islam, M. K., Biswas, C., Ahmed, H. U., Amin, M. A., et al. (2022). Exploration of EEG-Based depression biomarkers identification techniques and their applications: a systematic review. *IEEE Access* 10, 16756–16781. doi: 10.1109/ACCESS.2022.3146711

Fang, H., Tu, S., Sheng, J., and Shao, A. (2019). Depression in sleep disturbance: a review on a bidirectional relationship, mechanisms and treatment. *J. Cell Mol. Med.* 23, 2324–2332. doi: 10.1111/jcmm.14170 0

Freyburger, M., Pierre, A., Paquette, G., Bélanger-Nelson, E., Bedont, J., Gaudreault, P.-O., et al. (2016). EphA4 is involved in sleep regulation but not in the electrophysiological response to sleep deprivation. *Sleep* 39, 613–624. doi: 10.5665/sleep.5538

Garcia-Cortadella, R., Schwesig, G., Jeschke, C., Illa, X., Gray, A. L., Savage, S., et al. (2021). Graphene active sensor arrays for long-term and wireless mapping of wide frequency band epicortical brain activity. *Nat. Commun.* 12:211. doi: 10.1038/s41467-020-20546-w

Garn, H., Waser, M., Deistler, M., Benke, T., Dal-Bianco, P., Ransmayr, G., et al. (2015). Quantitative EEG markers relate to Alzheimer's disease severity in the Prospective Dementia Registry Austria (PRODEM). *Clin. Neurophysiol.* 126, 505–513. doi: 10.1016/j.clinph.2014.07.005

Gaubert, S., Raimondo, F., Houot, M., Corsi, M.-C., Naccache, L., Diego Sitt, J., et al. (2019). EEG evidence of compensatory mechanisms in preclinical Alzheimer's disease. *Brain* 142, 2096–2112. doi: 10.1093/brain/awz150

Ginzberg, R. (1955). Sleep and sleep disturbance in geriatric psychiatry. *J. Am. Geriatr. Soc.* 3, 493–511. doi: 10.1111/j.1532-5415.1955.tb00615.x

Goossens, J., Laton, J., Van Schependom, J., Gielen, J., Struyfs, H., Van Mossevelde, S., et al. (2017). EEG dominant frequency peak differentiates between Alzheimer's disease and frontotemporal lobar degeneration. *J. Alzheimers Dis.* 55, 53–58. doi: 10.3233/JAD-160188

Guarnieri, B., Adorni, F., Musicco, M., Appollonio, I., Bonanni, E., Caffarra, P., et al. (2012). Prevalence of sleep disturbances in mild cognitive impairment and dementing disorders: a multicenter Italian clinical cross-sectional study on 431 patients. *Dement. Geriatr. Cogn. Disord.* 33, 50–58. doi: 10.1159/000335363

Hamm, V., Héraud, C., Cassel, J.-C., Mathis, C., and Goutagny, R. (2015). Precocious alterations of brain oscillatory activity in Alzheimer's disease: a window of opportunity for early diagnosis and treatment. *Front. Cell. Neurosci.* 9:491. doi: 10.3389/fncel.2015.00491

Hata, M., Kazui, H., Tanaka, T., Ishii, R., Canuet, L., Pascual-Marqui, R. D., et al. (2016). Functional connectivity assessed by resting state EEG correlates with cognitive decline of Alzheimer's disease - an eLORETA study. *Clin. Neurophysiol.* 127, 1269–1278. doi: 10.1016/j.clinph.2015.10.030

Helfrich, R. F., and Knight, R. T. (2019). Cognitive neurophysiology: event-related potentials. *Handb. Clin. Neurol.* 160, 543–558.

Holz, J., Piosczyk, H., Feige, B., Spiegelhalder, K., Baglioni, C., Riemann, D., et al. (2012). EEG sigma and slow-wave activity during NREM sleep correlate with overnight declarative and procedural memory consolidation. *J. Sleep Res.* 21, 612–619. doi: 10.1111/j.1365-2869.2012.01017.x

Jeong, J. (2004). EEG dynamics in patients with Alzheimer's disease. *Clin. Neurophysiol.* 115, 1490–1505. doi: 10.1016/j.clinph.2004.01.001

Ju, Y.-E. S., Lucey, B. P., and Holtzman, D. M. (2014). Sleep and Alzheimer disease pathology–a bidirectional relationship. *Nat. Rev. Neurol.* 10, 115–119. doi: 10.1038/nrneurol.2013.269

Jung, F., and Carlén, M. (2021). "Chapter twelve - neuronal oscillations and the mouse prefrontal cortex," in *International Review of Neurobiology What does Medial Frontal Cortex Signal During Behavior? Insights from Behavioral Neurophysiology*, eds A. T. Brockett, L. M. Amarante, M. Laubach, and M. R. Roesch (Cambridge, MA: Academic Press), 337–372. doi: 10.1016/bs.irn.2020.11.005

Kaiser, A. K., Gnjezda, M.-T., Knasmüller, S., and Aichhorn, W. (2018). Electroencephalogram alpha asymmetry in patients with depressive disorders: current perspectives. *Neuropsychiatr. Dis. Treat* 14, 1493–1504. doi: 10.2147/NDT.S137776

Keizer, A. W. (2021). Standardization and personalized medicine using quantitative EEG in clinical settings. *Clin. EEG Neurosci.* 52, 82–89. doi: 10.1177/1550059419874945

Kent, B. A., Strittmatter, S. M., and Nygaard, H. B. (2018). Sleep and EEG power spectral analysis in three transgenic mouse models of Alzheimer's disease: APP/PS1, 3xTgAD, and Tg2576. *J. Alzheimers Dis.* 64, 1325–1336. doi: 10.3233/JAD-180260

Kropotov, J. D. (2016). "Chapter 2.3 - beta and gamma rhythms," in *Functional Neuromarkers for Psychiatry*, ed. J. D. Kropotov (San Diego, CA: Academic Press), 107–119. doi: 10.1016/B978-0-12-410513-3.00009-7

Kupfer, D. J., and Foster, F. G. (1972). Interval between onset of sleep and rapid-eye-movement sleep as an indicator of depression. *Lancet* 2, 684–686. doi: 10.1016/s0140-6736(72)92090-9

Liu, W., Miller, B. L., Kramer, J. H., Rankin, K., Wyss-Coray, C., Gearhart, R., et al. (2004). Behavioral disorders in the frontal and temporal variants of frontotemporal dementia. *Neurology* 62, 742–748. doi: 10.1212/01.wnl.0000113729.77161.c9

López-Sanz, D., Bruña, R., Delgado-Losada, M. L., López-Higes, R., Marcos-Dolado, A., Maestú, F., et al. (2019). Electrophysiological brain signatures for the classification of subjective cognitive decline: towards an individual detection in the preclinical stages of dementia. *Alzheimers Res. Ther.* 11:49. doi: 10.1186/s13195-019-0502-3

Matlis, S., Boric, K., Chu, C. J., and Kramer, M. A. (2015). Robust disruptions in electroencephalogram cortical oscillations and large-scale functional networks in autism. *BMC Neurol.* 15:97. doi: 10.1186/s12883-015-0355-8

McMackin, R., Bede, P., Pender, N., Hardiman, O., and Nasseroleslami, B. (2019). Neurophysiological markers of network dysfunction in neurodegenerative diseases. *Neuroimage Clin.* 22:101706. doi: 10.1016/j.nicl.2019.101706

McShane, B. B., Galante, R. J., Jensen, S. T., Naidoo, N., Pack, A. I., and Wyner, A. (2010). Characterization of the bout durations of sleep and wakefulness. *J. Neurosci. Methods* 193, 321–333. doi: 10.1016/j.jneumeth.2010.08.024

Medlej, Y., Asdikian, R., Wadi, L., Salah, H., Dosh, L., Hashash, R., et al. (2019). Enhanced setup for wired continuous long-term EEG monitoring in juvenile and adult rats: application for epilepsy and other disorders. *BMC Neurosci.* 20:8. doi: 10.1186/s12868-019-0490-z

Meghdadi, A. H., Stevanović Karić, M., McConnell, M., Rupp, G., Richard, C., Hamilton, J., et al. (2021). Resting state EEG biomarkers of cognitive decline associated with Alzheimer's disease and mild cognitive impairment. *PLoS One* 16:e0244180. doi: 10.1371/journal.pone.0244180

Minakawa, E. N., Wada, K., and Nagai, Y. (2019). Sleep disturbance as a potential modifiable risk factor for Alzheimer's disease. *Int. J. Mol. Sci.* 20:803. doi: 10.3390/ijms20040803

Morrone, C. D., Lai, A. Y., Bishay, J., Hill, M. E., and McLaurin, J. (2022). Parvalbumin neuroplasticity compensates for somatostatin impairment, maintaining cognitive function in Alzheimer's disease. *Trans. Neurodegenerat.* 11:26. doi: 10.1186/s40035-022-00300-6

Musiek, E. S., and Holtzman, D. M. (2016). Mechanisms linking circadian clocks, sleep, and neurodegeneration. *Science* 354, 1004–1008. doi: 10.1126/science.aah4968

Nakamura, A., Cuesta, P., Fernández, A., Arahata, Y., Iwata, K., Kuratsubo, I., et al. (2018). Electromagnetic signatures of the preclinical and prodromal stages of Alzheimer's disease. *Brain* 141, 1470–1485. doi: 10.1093/brain/awy044

Pavlenko, V. B., Chernyi, S. V., and Goubkina, D. G. (2009). EEG correlates of anxiety and emotional stability in adult healthy subjects. *Neurophysiology* 41, 337–345. doi: 10.1007/s11062-010-9111-2

Pinnell, R. C., Almajidy, R. K., Kirch, R. D., Cassel, J. C., and Hofmann, U. G. (2016). A wireless EEG recording method for rat use inside the water maze. *PLoS One* 11:e0147730. doi: 10.1371/journal.pone.0147730

Pinnell, R. C., Dempster, J., and Pratt, J. (2015). Miniature wireless recording and stimulation system for rodent behavioural testing. *J. Neural Eng.* 12:066015. doi: 10.1088/1741-2560/12/6/066015

Saito, T., Matsuba, Y., Mihira, N., Takano, J., Nilsson, P., Itohara, S., et al. (2014). Single app knock-in mouse models of Alzheimer's disease. *Nat. Neurosci.* 17, 661–663. doi: 10.1038/nn.3697

Saxena, S., and Liebscher, S. (2020). Editorial: circuit mechanisms of neurodegenerative diseases. *Front. Neurosci.* 14:593329. doi: 10.3389/fnins.2020.593329

Scullin, M. K., and Bliwise, D. L. (2015). Sleep, cognition, and normal aging: integrating a half century of multidisciplinary research. *Perspect. Psychol. Sci.* 10, 97–137. doi: 10.1177/1745691614556680

Snyder, S. M., Rugino, T. A., Hornig, M., and Stein, M. A. (2015). Integration of an EEG biomarker with a clinician's ADHD evaluation. *Brain Behav.* 5:e00330. doi: 10.1002/brb3.330

Tononi, G., and Cirelli, C. (2014). Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron* 81, 12–34. doi: 10.1016/j.neuron.2013.12.025

Tort, A. B. L., Komorowski, R. W., Manns, J. R., Kopell, N. J., and Eichenbaum, H. (2009). Theta-gamma coupling increases during the learning of item-context associations. *Proc. Natl. Acad. Sci. U S A.* 106, 20942–20947. doi: 10.1073/pnas.0911331106

van der Zande, J. J., Gouw, A. A., van Steenoven, I., van de Beek, M., Scheltens, P., Stam, C. J., et al. (2020). Diagnostic and prognostic value of EEG in prodromal dementia with Lewy bodies. *Neurology* 95, e662–e670. doi: 10.1212/WNL.0000000000009977

Vitiello, M. V., and Borson, S. (2001). Sleep disturbances in patients with Alzheimer's disease: epidemiology, pathophysiology and treatment. *CNS Drugs* 15, 777–796. doi: 10.2165/00023210-200115100-00004

Vitiello, M. V., Poceta, J. S., and Prinz, P. N. (1991). Sleep in Alzheimer's disease and other dementing disorders. *Can. J. Psychol.* 45, 221–239. doi: 10.1037/h0084283

Vogler, E. C., Flynn, D. T., Busciglio, F., Bohannan, R. C., Tran, A., Mahavongtrakul, M., et al. (2017). Low cost electrode assembly for EEG recordings in mice. *Front. Neurosci.* 11:629. doi: 10.3389/fnins.2017.00629

Vorobyov, V., Bakharev, B., Medvinskaya, N., Nesterova, I., Samokhin, A., Deev, A., et al. (2019). Loss of midbrain dopamine neurons and altered apomorphine EEG effects in the 5xFAD mouse model of Alzheimer's disease. *J. Alzheimers Dis.* 70, 241–256. doi: 10.3233/JAD-181246

Wainberg, M., Jones, S. E., Beaupre, L. M., Hill, S. L., Felsky, D., Rivas, M. A., et al. (2021). Association of accelerometer-derived sleep measures with lifetime psychiatric diagnoses: a cross-sectional study of 89,205 participants from the UK Biobank. *PLoS Med.* 18:e1003782. doi: 10.1371/journal.pmed.1003782

Waite, F., Sheaves, B., Isham, L., Reeve, S., and Freeman, D. (2020). Sleep and schizophrenia: from epiphenomenon to treatable causal target. *Schizophr. Res.* 221, 44–56. doi: 10.1016/j.schres.2019.11.014

Walker, M. P., and Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron* 44, 121–133. doi: 10.1016/j.neuron.2004.08.031

Wang, J., Barstein, J., Ethridge, L. E., Mosconi, M. W., Takarae, Y., and Sweeney, J. A. (2013). Resting state EEG abnormalities in autism spectrum disorders. *J. Neurodev. Disord.* 5:24. doi: 10.1186/1866-1955-5-24

Wang, Y.-Q., Li, R., Zhang, M.-Q., Zhang, Z., Qu, W.-M., and Huang, Z.-L. (2015). The neurobiological mechanisms and treatments of REM sleep disturbances in depression. *Curr. Neuropharmacol.* 13, 543–553. doi: 10.2174/1570159x13666150310002540

Webb, C. A., Cui, R., Titus, C., Fiske, A., and Nadorff, M. R. (2018). Sleep disturbance, activities of daily living, and depressive symptoms among older adults. *Clin. Gerontol.* 41, 172–180. doi: 10.1080/07317115.2017.1408733

Xhima, K., Markham-Coultes, K., Hahn Kofoed, R., Saragovi, H. U., Hynynen, K., and Aubert, I. (2022). Ultrasound delivery of a TrkA agonist confers neuroprotection to Alzheimer-associated pathologies. *Brain* 145, 2806–2822. doi: 10.1093/brain/awab460

Xu, W., Fitzgerald, S., Nixon, R. A., Levy, E., and Wilson, D. A. (2015). Early hyperactivity in lateral entorhinal cortex is associated with elevated levels of AβPP metabolites in the Tg2576 mouse model of Alzheimer's disease. *Exp. Neurol.* 264, 82–91. doi: 10.1016/j.expneurol.2014.12.008

Zayachkivsky, A., Lehmkuhle, M. J., and Dudek, F. E. (2015). Long-term continuous EEG monitoring in small rodent models of human disease using the epoch wireless transmitter system. *J. Vis. Exp.* e52554. doi: 10.3791/52554

Glossary

Barnes maze: a cognitive task for spatial learning and memory in mice; Bout: a period of time spent asleep or awake, defined herein as three continuous 10-second epochs; EDF: European Data Format; a standard time series file format commonly used for wave-form data such as EEG and EMG; Epoch: period of time of wave-form data utilized to stage sleep (10-seconds herein); EthoVision XT: Noldus behavior acquisition and analysis software; Non-rapid eye movement (NREM) sleep: sleep stages including restorative slow wave sleep, characterized by delta waves in EEG; PhenoTyper: a home-cage designed for continuous video assessment of mice; Potentiostat: recording device connected to the headcaps to wirelessly acquire and transmit EEG and EMG data; Rapid eye movement (REM) sleep: sleep stage involved in memory consolidation and dreaming, distinguished from NREM by increased theta and reduced delta waves in EEG; Sirenia®: Pinnacle acquisition and analysis software for EEG and EMG data; Transistor-transistor logic (TTL): for circuit integration, can be used to align electrical signal within multiple software; tsv: tab-separated values text file.

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 23 January 2023
DOI 10.3389/fnagi.2022.899249

OPEN ACCESS

EDITED BY
Yang Jiang,
University of Kentucky,
United States

REVIEWED BY
Nadia Solaro,
University of Milano-Bicocca, Italy
Edward Lakatta,
National Institute on Aging (NIH),
United States
Frederic Roche,
Université Jean Monnet,
France
Chih-Cheng Huang,
Kaohsiung Chang Gung Memorial Hospital,
Taiwan

*CORRESPONDENCE
Andy Schumann
✉ andy.schumann@med.uni-jena.de

SPECIALTY SECTION
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 18 March 2022
ACCEPTED 20 December 2022
PUBLISHED 23 January 2023

CITATION
Schumann A, Gaser C, Sabeghi R, Schulze PC, Festag S, Spreckelsen C and Bär K-J (2023) Using machine learning to estimate the calendar age based on autonomic cardiovascular function. *Front. Aging Neurosci.* 14:899249. doi: 10.3389/fnagi.2022.899249

Using machine learning to estimate the calendar age based on autonomic cardiovascular function

Andy Schumann[1]*, Christian Gaser[2,3], Rassoul Sabeghi[1], P. Christian Schulze[4], Sven Festag[5,6], Cord Spreckelsen[5,6] and Karl-Jürgen Bär[1]

[1]Lab for Autonomic Neuroscience, Imaging and Cognition (LANIC), Department of Psychosomatic Medicine and Psychotherapy, Jena University Hospital, Jena, Germany, [2]Hans Berger Department of Neurology, Jena University Hospital, Jena, Germany, [3]Department of Psychiatry and Psychotherapy, Jena University Hospital, Jena, Germany, [4]Department of Internal Medicine I, Division of Cardiology, Jena University Hospital, Jena, Germany, [5]Institute of Medical Statistics, Computer and Data Sciences, Jena University Hospital, Jena, Germany, [6]SMITH Consortium of the German Medical Informatics Initiative, Leipzig, Germany

Introduction: Aging is accompanied by physiological changes in cardiovascular regulation that can be evaluated using a variety of metrics. In this study, we employ machine learning on autonomic cardiovascular indices in order to estimate participants' age.

Methods: We analyzed a database including resting state electrocardiogram and continuous blood pressure recordings of healthy volunteers. A total of 884 data sets met the inclusion criteria. Data of 72 other participants with an BMI indicating obesity (>30 kg/m^2) were withheld as an evaluation sample. For all participants, 29 different cardiovascular indices were calculated including heart rate variability, blood pressure variability, baroreflex function, pulse wave dynamics, and QT interval characteristics. Based on cardiovascular indices, sex and device, four different approaches were applied in order to estimate the calendar age of healthy subjects, i.e., relevance vector regression (RVR), Gaussian process regression (GPR), support vector regression (SVR), and linear regression (LR). To estimate age in the obese group, we drew normal-weight controls from the large sample to build a training set and a validation set that had an age distribution similar to the obesity test sample.

Results: In a five-fold cross validation scheme, we found the GPR model to be suited best to estimate calendar age, with a correlation of r=0.81 and a mean absolute error of MAE=5.6 years. In men, the error (MAE=5.4 years) seemed to be lower than that in women (MAE=6.0 years). In comparison to normal-weight subjects, GPR and SVR significantly overestimated the age of obese participants compared with controls. The highest age gap indicated advanced cardiovascular aging by 5.7 years in obese participants.

Discussion: In conclusion, machine learning can be used to estimate age on cardiovascular function in a healthy population when considering previous models of biological aging. The estimated age might serve as a comprehensive and readily interpretable marker of cardiovascular function. Whether it is a useful risk predictor should be investigated in future studies.

KEYWORDS

aging, heart rate variability, blood pressure variability, baroreflex, pulse pressure

Introduction

Maintaining a healthy cardiovascular system is one of the most important goals of modern health policy (Mendis et al., 2011). Factors elevating cardiovascular risk include physical inactivity and an unhealthy diet. In addition, age is an independent risk factor for the development of cardiovascular disease (CVD)—the leading cause of death worldwide.

The cardiovascular system is a complex structure that comprises the heart and vasculature that are not under voluntary control. Instead, the autonomic nervous system adapts the activity of the heart and vascular tone to changing environmental demands. To assess the state of the cardiovascular system, physicians usually estimate blood pressure and record electrocardiograms (ECGs). Several indicators of cardiovascular risk can be determined from these data.

Heart rate and its variability (HRV) are established markers of cardiac fitness (Jensen et al., 2013; Nanchen et al., 2013). A natural decay in HRV during the course of aging is a consistent finding of several studies (Reardon and Malik, 1996; Fukusaki et al., 2000; Boettger et al., 2010; Voss et al., 2012, 2015). Lower levels of HRV have been associated with increased cardiovascular morbidity and mortality in the elderly (Tsuji et al., 1996). The feedback loop that adapts heart rate to changes in blood pressure, that is, baroreflex function, is progressively diminished with increasing age (Laitinen et al., 2004). Various factors such as endothelial dysfunction or oxidative stress result in the stiffening of large arteries, which is a condition promoting sustained hypertension, atherosclerosis, and thrombosis (Dai et al., 2015). Indicators of age-related vascular changes are broad pulse waves, elevated pulse wave velocity, and increased systolic blood pressure. Considering the widespread effects of aging on the cardiovascular system, it seems useful to combine different established indices into one comprehensive marker of cardiovascular health.

Estimating age based on biological data is a widely used concept in other medical disciplines, for instance, to evaluate brain health (Gialluisi et al., 2019). Aging affects different aspects of brain structure and function that can be summarized as the estimated age of the brain (Franke and Ten Gaser, 2019; Dafflon et al., 2020; Jiang et al., 2020). Using this framework, scientists were able to trace brain development and to assess the risk of developing neurodegenerative diseases and general mortality in older adults (see Cole and Franke, 2017).

Recently, machine learning (ML) methods have gained a lot of attention in efforts to improve risk prediction and clinical outcomes in patients with cardiovascular (see Sevakula et al., 2020, for review). ML algorithms can be used to automatically identify information that will help solve a given problem. Supervised learning methods build an analytical model based on a set of training samples containing input and related output values. Applying this model to a test set of input data without knowing the desired output reveals the accuracy of the automatic solution. For regression problems, an output function is obtained by fitting a line to the data points in a high-dimensional space built from available input variables (feature space; Bennett and Campbell, 2000; Schölkopf and Smola, 2002). Assessing cardiovascular risk by ML has been demonstrated to be more accurate than conventional approaches (Kakadiaris et al., 2018) with a lower bias than non-ML methods (Suri et al., 2022).

In this study, we aimed to estimate age based on cardiovascular data by applying ML. Input features were extracted from simultaneous resting recordings of ECG and continuous blood pressure in healthy individuals. We compared different approaches to solve regression problems, namely, support vector regression, relevance vector regression, Gaussian process regression, and a linear regression model. In a proof-of-concept application, we compared age estimates in obese but otherwise healthy individuals and normal-weight controls. As obesity is related to an impairment of cardiovascular function and elevated cardiovascular risk, we assumed systematically higher age estimates when compared to normal-weight controls. Thus, we derived three age-matched subsamples from our database. We trained all ML models on normal-weight controls and applied them to a sample of obese individuals and an independent sample of normal-weight controls.

Materials and methods

Database

Resting-state physiological recordings of 1,121 healthy volunteers were obtained. None of the subjects had any history of neurological or psychiatric disorders. Exclusion criteria were any medical conditions, illegal drugs, or medication potentially influencing cardiovascular function. Thorough physical examination, resting electrocardiography (ECG), and routine laboratory parameters (electrolytes, basic metabolic panel, and blood count) had to be without any pathological finding. All participants provided written informed consent before participating in the study. The study protocol was approved by the Ethics Committee of the University Hospital of Jena (#5423-01/18, 4,940-10/16). Data sets have been made publicly available at *PhysioNet* (Goldberger et al., 2000; Schumann and Bär, 2022).

From the original database, 31 subjects were excluded due to missing or invalid information regarding age, gender, or BMI. A total of 118 recordings were excluded due to incomplete or missing blood pressure recordings. A set comprising data from 72 obese participants (body mass index (BMI) > 30) was excluded from the main set but used for additional assessments during a later stage. Cardiovascular indices were estimated for the 884 healthy subjects from the main set (59% females, age: 31.3 ± 13.6 years, BMI: 22.8 ± 2.8 kg/m^2, see Table 1) and for the 72 obese participants.

Data recordings

Continuous non-invasive blood pressure and ECG were acquired simultaneously over 20 min in the supine position using either a Task Force Monitor® (TFM, CNSystems Medizintechnik GmbH, Graz, Austria) or MP150 (BIOPAC Systems Inc., Goleta, CA, United States). First 5 min were excluded from the analysis. R-waves and systolic and diastolic blood pressure values were extracted from the data using automatic detection algorithms delivered with the devices (Task Force® Monitor, CNSystems or AcqKnowledge 4.1, BIOPAC Systems). An adaptive filter procedure was applied to identify and substitute premature ventricular beats and artifacts based on the heart beat intervals (Wessel et al., 2000). Data sets with an artifact rate larger than 5% of all intervals were excluded from the analysis.

Estimation of cardiovascular indices

From the ECG-derived heart beat interval time series (BBI), we calculated the mean heart rate (HR), root-mean-square of successive BBI (RMSSD), the standard deviation of BBI (SDNN), low- and high-frequency power and their ratio (Malik et al., 1996), deceleration capacity (Bauer et al., 2006), Renyi entropy (base 1/4; Renyi, 1961), sample entropy (Richman and Moorman, 2000), and compression entropy (Baumert et al., 2004). The mean and standard deviation of corrected QT intervals (Hodges et al., 1983) and the QT variability index (QTVI) were estimated (Berger, 2003).

TABLE 1 Sample characteristics separated by recording device.

	TFM	MP150
Age [y]	31.9 ± 14.3	30.5 ± 12.6
N	18–92	18–82
Sex (f/m)	283/205	240/156
BMI [kg/m^2]	22.7 ± 2.9	22.8 ± 2.7

Age and BMI are given in mean value ± standard deviation. Data have been recorded using two different devices (see the Data Recordings section, for more details). TFM, Recorded by CNSystems Task Force Monitor; MP150, Recorded by BIOPAC Systems MP150; BMI, body mass index.

From continuous blood pressure, the mean and standard deviation of systolic blood pressure (SBP) and diastolic blood pressure (DBP) values per heart beat interval were extracted (Floras, 2013). Pulse pressure was calculated as differences between SBP and DBP. Using the dual sequence method, baroreflex sensitivity was calculated as a marker of bradycardic and tachycardic changes due to blood pressure alterations (Malberg et al., 1999). Mean values and standard deviation of the pulse transit time, pulse rise time, pulse wave duration, pulse wave velocity, and time delay of the dicrotic notch were estimated on blood pressure signals (Fischer et al., 2017; Table 2).

Age estimation

Four different ML approaches were applied to estimate calendar age based on 29 cardiovascular indices and the two categorial variables sex and recording device. The algorithms have been implemented in Python version 3.8.3 using the toolbox scikit-learn version 0.24.1 (Pedregosa et al., 2011).

Gaussian process regression (GPR) models use a kernel to define the covariance of a distribution over the target functions and observed training data to define a likelihood function (Schulz et al., 2018). We used a combination of a constant kernel with a radial basis function (RBF). **Support vector regression (SVR)** models offer the flexibility to define how much error is acceptable in finding an appropriate fit to the input data (Vapnik, 1995; Schölkopf and Smola, 2002). An RBF kernel and regularization index C = 30 were used. **Relevance vector regression (RVR)** models use Bayesian inference to obtain parsimonious solutions for regression (Tipping, 2000). Here, we also used an RBF kernel. Hyperparameters of GPR, SVR, and RVR estimation were optimized using grid search. The performance of these approaches was compared to a **linear regression (LR)** model that estimates age as a linear combination of the input variables.

Model performance

In a 5-fold cross-validation scheme, one-fifth of the main set was randomly assigned to a test set, while the model was trained on the remaining four-fifths of the data. In each of the five runs, another fifth of the data served as test data. After the five runs, the empiric means (± standard error) of the evaluation metrics are reported as the final ones. The cross-validation was repeated 20 times with a randomized order of input data. Again, the metrics were averaged over all repetitions. We standardized all input data before using training and test data during the cross-validation procedure (*StandardScaler* implemented in *sklearn*).

TABLE 2 Indices included in age estimation.

Index	Explanation
Standard heart rate variability (HRV)	
HR [min^{-1}]	Mean heart rate
SDNN [ms]	Standard deviation of heart beat intervals (BBI)
RMSSD [ms]	Root-mean-square of successive BBI differences
DC [ms]	Deceleration capacity
Spectral HRV	
LF [ms^2]	Low frequency spectral power of BBI
HF [ms^2]	High-frequency spectral power of BBI
LF/HF [a.u.]	Low-to-high frequency spectral power ratio
Nonlinear HRV	
CompEn [a.u.]	Compression entropy
SampEn [a.u.]	Sample entropy
RenyiEn [bit]	Renyi entropy
Cardiovascular regulation	
BRS [ms/mmHg]	Baroreflex sensitivity
LFalpha [ms/mmHg]	Low frequency cardiovascular coherence
HFalpha [ms/mmHg]	High-frequency cardiovascular coherence
JSDsym [a.u.]	Symmetric joint symbolic dynamics
SBP [mmHg]	Mean systolic blood pressure
sd_SBP [mmHg]	Standard deviation of systolic blood pressure
DBP [mmHg]	Mean diastolic blood pressure
sd_DBP [mmHg]	Standard deviation of diastolic blood pressure
PP [mmHg]	Pulse pressure
Pulse wave dynamics	
PTT [ms]	Mean pulse transit time
sdPTT [ms]	Standard deviation of PTT
PRT [ms]	Mean pulse rise time
sdPRT [ms]	Standard deviation of PRT
PWV [mmHg/ms]	Mean pulse wave velocity
sdPWV [mmHg/ms]	Standard deviation of PWV
SIT [ms]	Mean delay of dicrotic notch to pulse maximum
QT interval characteristics	
meanQTc [ms]	Mean corrected QT interval
sdQTc [ms]	Standard deviation of QT interval
QTVI [a.u.]	QT variability index
Additional info	
Sex [0/1]	Participant gender
Device [0/1]	Recording device

The quality of age estimation was evaluated by the mean absolute error (MAE), root-mean-squared error (RMSE), and Pearson's correlation (r) of estimated $\hat{Y}_i$ and the actual age y_i.

$$MAE = \frac{\sum_i \left(\left| \hat{Y}_i - Y_i \right| \right)}{N}$$

$$RMSE = \sqrt{\frac{\sum_i \left(\hat{Y}_i - Y_i \right)^2}{N}}$$

Comparison of weight groups

In a second experiment, we aimed to compare age estimates in obese participants (BMI > 30 kg/m^2) with normal-weight peers. Therefore, we draw two subsamples from our normal-weight population that served as training and test data. The calendar age of these data should match our obese sample. Therefore, we categorized participants into age groups of 10 years (see Table 3) and estimated the relative distribution of obese participants across these age groups. Then, we randomly assigned normal-weight participants from each age group to training and a test set to match the age distribution of the obese sample. Finally, we trained all ML models on 197 normal-weight individuals (101 women, 96 men, 41.4 ± 14.9 years, BMI: 23.4 ± 2.2 kg/m^2) to estimate age in the obese test sample of 72 individuals (44 women, 28 men, age: 42.9 ± 15.5 years, BMI: 34.8 ± 5.8 kg/m^2) in a normal test sample of 72 normal-weight controls (37 women, 35 men, age: 42.1 ± 15.8 years, BMI: 23.3 ± 2.4 kg/m^2). We used mean values and standard deviation of training data to standardize the training, normal and obese test sets (*StandardScaler*, *sklearn*). The age gap (deviation between estimated and calendar age) was calculated and compared between the normal-weight and the obese test set using the one-sided Wilcoxon rank-sum test.

Results

The final sample under investigation included 884 healthy individuals. In Table 3, sample characteristics are depicted within different age ranges.

A number of relevant autonomic cardiovascular indices are depicted in Figure 1. It seems obvious from these plots that age has a different effect on each of those measures. For instance, systolic blood pressure seems to increase almost linearly with age, while HRV decreases with age rather exponentially. In total, 29 different indices together with *sex* and *device* served as input features for age prediction models.

In Figure 2, we plotted the age estimated by Gaussian process regression (GPR) against the calendar age for one cross-validation run (Figure 2; r = 0.81, MAE = 5.62 years, RMSE = 8.00 years). In this scatter plot, it becomes clear that most of the data sets are in the lower age range. At a higher age (over 70 years), the model tends to underestimate the individual age.

Comparison of methods

After 20 runs of the 5-fold cross-validation, we compared performance measures of age prediction by four different mathematical models (see Table 4). Lowest error estimates were achieved using GPR to estimate age (MAE = 5.6 years and RMSE = 8.0 years). The highest correlation was also calculated between GPR estimates and underlying calendar age. Overall performance tended to be better in men compared to women. However, in women and men, GPR was the most accurate approach to estimate age in terms of errors and correlation to underlying calendar age.

TABLE 3 Sample description in different age ranges.

	Healthy participants in age groups				
	<30 years	30–39 years	40–49 years	50–59 years	≥60 years
Age [y]	23.7 ± 2.7	33.6 ± 2.9	44.4 ± 2.8	53.5 ± 2.7	70.1 ± 9
N	585	123	75	47	54
Sex (f/m)	386/199	54/69	30/45	16/31	38/16
BMI [kg/m^2]	22.1 ± 2.6	23.2 ± 2.7	24.3 ± 2.9	25.2 ± 2.7	24.8 ± 2.5

Results are given in mean values ± standard deviation. BMI: body mass index.

Age gap estimation

Figure 3 shows the estimated autonomic indices for the normal-weight train group (normal train), the normal-weight test group (normal test), and the obese test group (obese test). Results from the group comparison between the two test sets are shown in Supplementary Table S1. Compared to normal participants, obese individuals showed alterations in a number of autonomic indices, including elevated heart rates, reduced vagal heart rate variability (RMSSD) and baroreflex sensitivity, and increased blood pressure (see Figure 3). The two normal-weight groups did not seem to differ significantly from one another except for the probability of symmetric symbolic dynamics of blood pressure and heart rate (JSDsym, see Supplementary Table S1).

All four methods were used to estimate the age gap between normal-weight healthy individuals ($N = 72$) and obese but apparently healthy individuals ($N = 72$). In Figure 4, deviations between calendar age and ML estimates per method are illustrated. The average errors of age estimation were higher in obese participants for all four methods. The Wilcoxon rank-sum

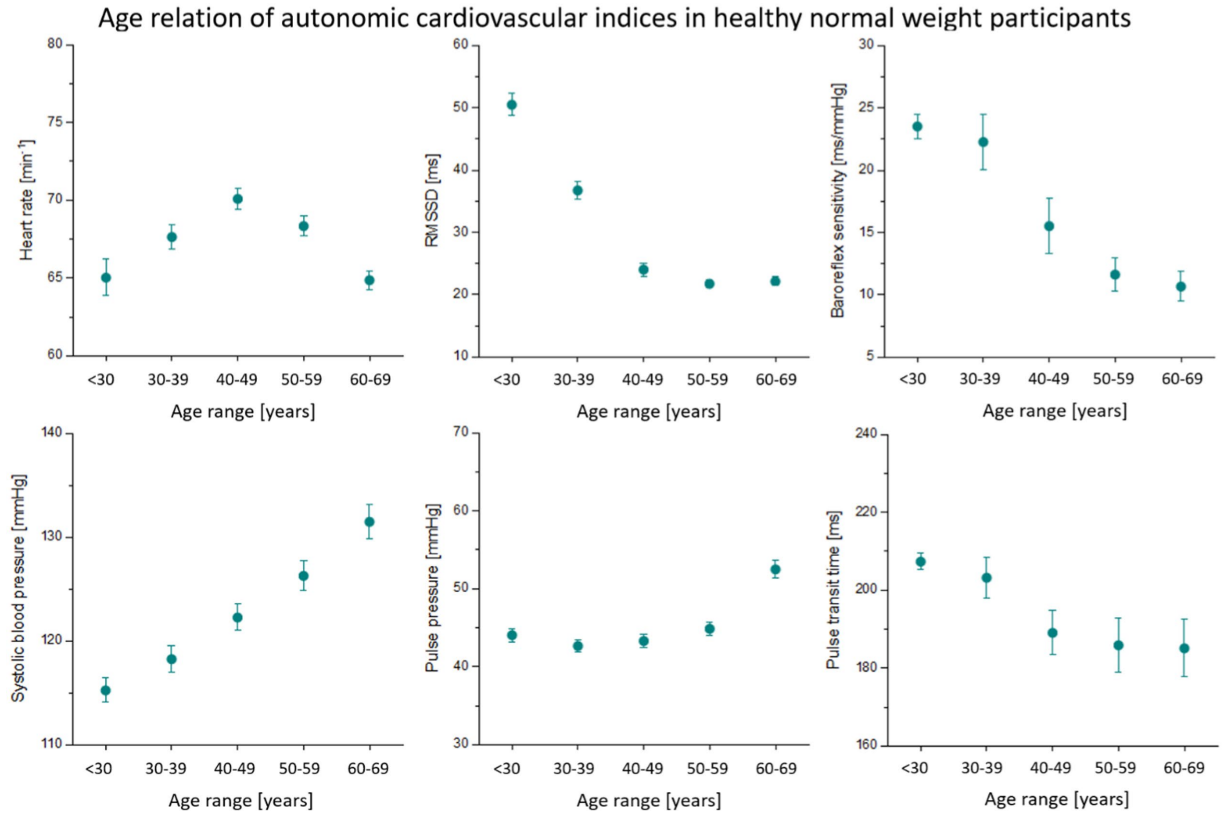


FIGURE 1
Age dependency of autonomic cardiovascular indices. Mean values and 95% confidence intervals are depicted. RMSSD, vagal heart rate variability (root-mean-square of successive heart beat intervals).

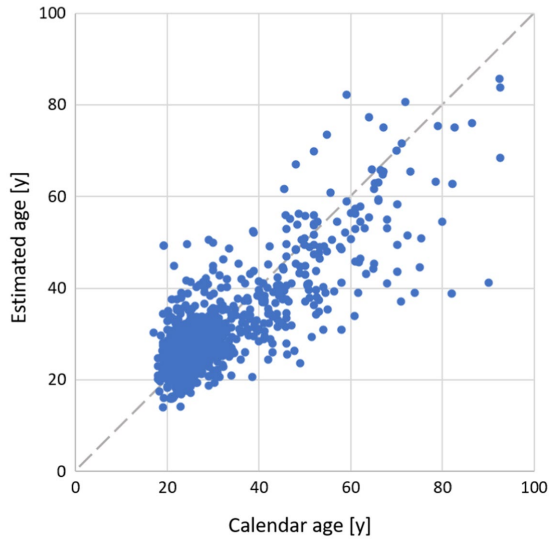


FIGURE 2
Relation of individual age estimated by GPR and calendar age. Dashed gray line indicates perfect concordance.

TABLE 4 Performance scores of age prediction after 20 repetitions of cross-validation.

Performance scores	Regression approach			
	LR	GPR	SVR	RVR
MAE [y]	6.39 ± 0.36	**5.64 ± 0.33**	6.45 ± 0.33	6.15 ± 0.8
RMSE [y]	8.67 ± 0.06	**8.03 ± 0.08**	8.86 ± 0.13	8.71 ± 0.27
r	0.77 ± 0.01	**0.81 ± 0.01**	0.77 ± 0.01	0.77 ± 0.01
Performance scores in women (N = 523)				
MAE [y]	6.47 ± 0.06	**6.01 ± 0.1**	6.96 ± 0.15	6.5 ± 0.3
RMSE [y]	8.53 ± 0.09	**8.37 ± 0.13**	9.48 ± 0.2	9.06 ± 0.37
r	0.73 ± 0.01	**0.74 ± 0.01**	0.68 ± 0.01	0.7 ± 0.02
Performance scores in men (N = 361)				
MAE [y]	6.33 ± 0.05	**5.38 ± 0.06**	6.09 ± 0.1	5.91 ± 0.23
RMSE [y]	8.76 ± 0.08	**7.79 ± 0.09**	8.4 ± 0.13	8.46 ± 0.27
r	0.79 ± 0.01	**0.83 ± 0.01**	0.81 ± 0.01	0.8 ± 0.01

The lowest error estimates and highest correlation coefficient are written in bold. LR, linear regression; GPR, Gaussian process regression; SVR, support vector regression; RVR, relevance vector regression; r, Pearson's correlation coefficient; MAE, mean absolute error; RMSE, root-mean-squared error.

test revealed a significantly increased age estimation error in obese participants by 5.4 years (interquartile range IQR = [−7.2; 13.1]; z = −1.744, p = 0.0406) when using GPR and 5.7 years (IQR = [−7.2;13.6]; z = −2.148; p = 0.0159) when using SVR.

Discussion

As artificial intelligence has entered most aspects of our daily life, it is not surprising that machine learning (ML) is about to revolutionize the medical sector and healthcare industry (Chen et al., 2017; Koch, 2018). ML offers great opportunities to improve risk stratification, diagnostic classification, clustering for the identification of patient subgroups, and many more (Sevakula et al., 2020). One popular application is the quantification of aging effects based on biological information (Al Zoubi et al., 2018; Dafflon et al., 2020; Jiang et al., 2020). To the best of our knowledge, this is the first study to use autonomic markers from cardiovascular recordings to predict chronological age in healthy subjects using ML.

Estimated age was strongly correlated with actual age with an error of MAE = 5.6 years, RMSE = 8.0 years, and r = 0.81 when GPR was used. However, the accuracy of age estimation was lower in older participants. For instance, age was underestimated in participants over 70 years of age by the GPR model (Figure 2). As data sets used in our study are concentrated at a younger age (<40 years), models are primarily trained on young individuals. This makes the prediction of age more difficult in the elderly. The accuracy of the models was higher in men than in women. This might be due to the fact that some age-related changes are more pronounced in men than in women. Especially, blood pressure and vascular indicators of arterial stiffening have been reported to correlate stronger with age in men (AlGhatrif et al., 2013). Maybe an additional influence of the menstrual cycle may have increased the variance of estimated features in women that we were not able to account for (Schmalenberger et al., 2020).

Considering state-of-the-art approaches to estimate age based on biological information, as reviewed by Gialluisi et al. (2019), the accuracy of our models was quite high. According to their summary, brain data and blood markers have been more widely used to estimate age, with MAEs ranging between 4.2 and 11.8 years. Only five of the 14 studies that have been reviewed achieved MAEs below 5 years—all based on brain data. Analyzing blood values, the most accurate model had an MAE of 5.6 years. In contrast, ML approaches to evaluate the aging of the cardiovascular system are rare. Using linear regression models, few studies have already attempted to estimate age based on HRV and ECG. Colosimo (1997) used a linear model to estimate age that correlated with calendar age with r = 0.71 in 141 subjects. More recently, Starc et al. (2012) predicted age using HRV and a multiple linear regression model with a high correlation of r = 0.87 in 377 subjects. Unlike those approaches, ML techniques automatically determine a numerical solution from a variety of input data through the learning process. Input features can be of different types (scalar values, signals, and images), and finally, they can contribute in a nonlinear fashion to this solution. Therefore, ML strategies often improve the accuracy of mathematical models in several applications (e.g., Acevedo et al., 2009; Ren et al., 2020).

In this study, we used a variety of established cardiovascular indices as input features. However, there are countless measures of heart rate variability alone. Instead of calculating these variables on physiological recordings, the recorded signals themselves can be entered into the models. Relevant signal segments then contribute to the estimation of age, avoiding the selection of suitable cardiovascular indices. *Via* methods of deep learning,

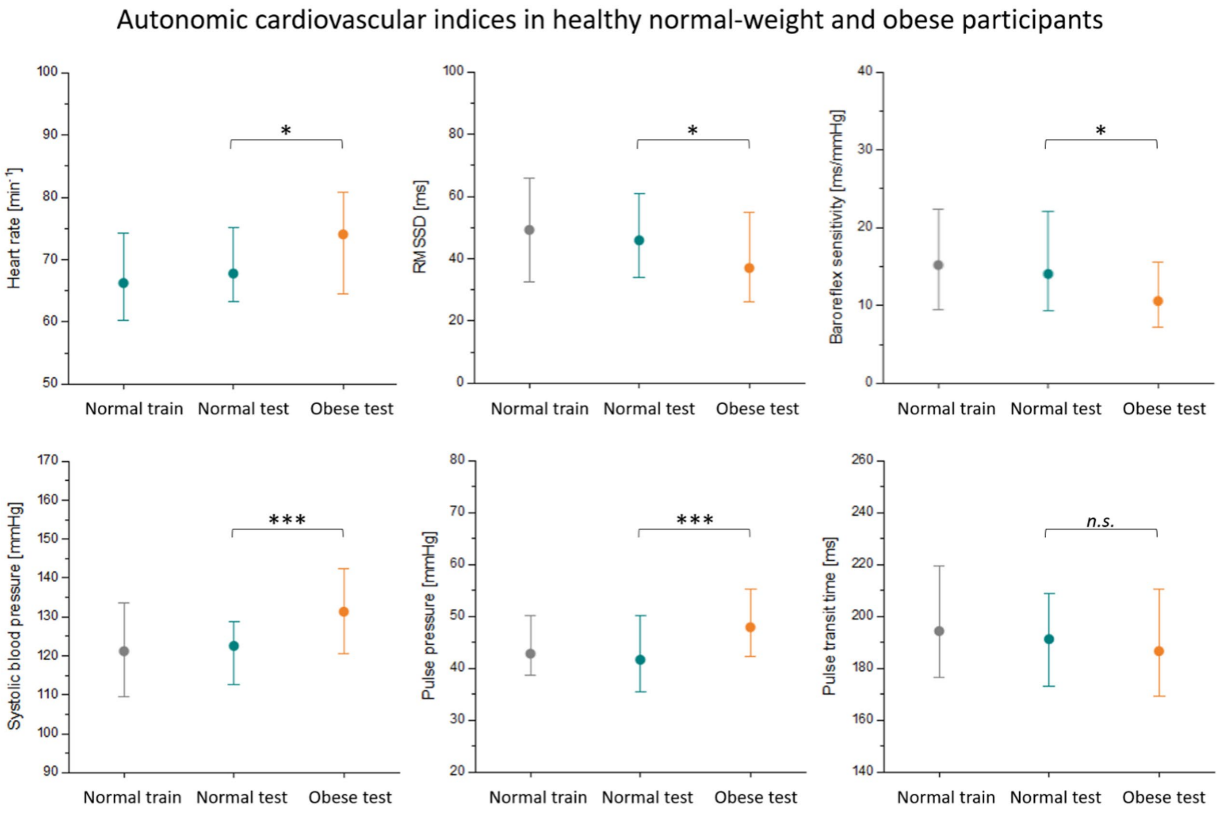


FIGURE 3
Autonomic cardiovascular indices in two subsets of normal-weight participants (training data in gray, test data in cyan) and obese participants (orange). The median is depicted together with the lower quartile and the upper quartile. The two test sets were compared using the Wilcoxon rank-sum test with *p*-values indicated in the figure [$p<0.05$ (*); $p<0.001$ (***); $p>0.05$ (n.s.)]. RMSSD, vagal heart rate variability (root-mean-square of successive heart beat intervals).

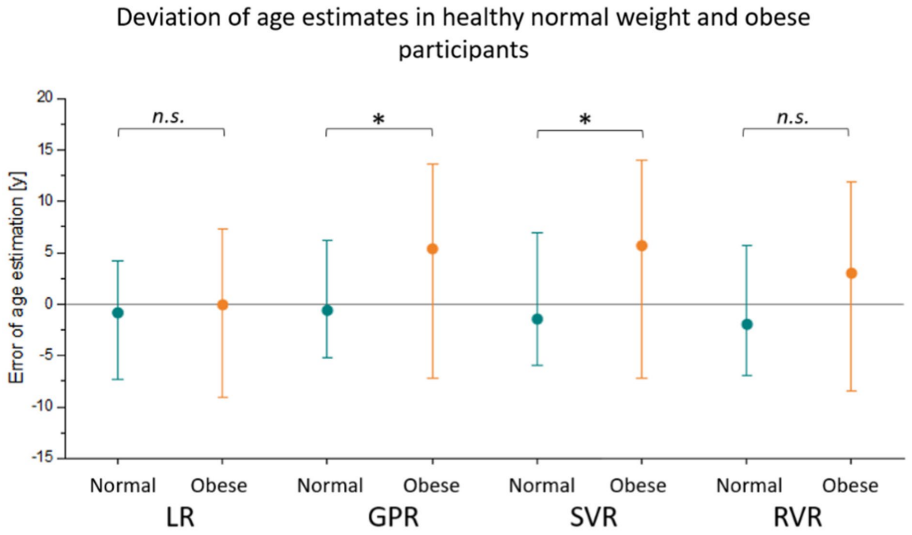


FIGURE 4
Deviation of estimated age from calendar age in a sample of obese participants (orange) and an independent set of matched normal-weight controls (cyan). The median is depicted together with the lower quartile and the upper quartile. Four different models, namely linear regression (LR), Gaussian process regression (GPR), support vector regression (SVR), and relevance vector regression (RVR), were trained on matched normal-weight individuals. The age estimation errors were compared between the normal-weight and obese test set using the Wilcoxon rank-sum test with p-values indicated in the figure [$p<0.05$ (*); $p>0.05$ (n.s.)].

Attia et al. (2019) estimated the calendar age from short 12-lead ECG signals. A convolutional neural network led to an error of MAE = 6.9 years and r = 0.84. Those patients, whose predicted age was more than 7 years higher than their calendar age, were more likely to be diagnosed with cardiovascular diseases, such as hypertension or coronary disease. The authors acknowledged that one key limitation to the findings in their study was the fact that the large underlying sample of 774,783 subjects included only patients who had their ECG recorded for some clinical indication. Similarly, Strodthoff et al. (2021) estimated calendar age based on short 12-lead ECG records using different neural networks. For age estimation, a feedforward residual neural network performed best with an error of MAE = 6.86 years and r = 0.85. Their database included a total of 21,837 both normal and abnormal clinical ECG recordings of 18,885 patients.

In our study, we estimated the deviation from normal healthy aging in a sample of obese but otherwise healthy participants as proof-of-concept. All four models were trained on normal-weight controls and then used to estimate age in a normal-weight and an obese test set. Calendar age distribution was matched across subsets. Using GPR and SVR, the age gap between estimated age and calendar age was significantly higher in obese participants than in normal-weight controls. This means that there was a systematic overestimation of age in obese participants. At least some differences in cardiovascular indices between obese and normal-weight participants are similar to changes that occur during aging. We observed elevated systolic blood pressure and pulse pressure as well as lower vagal HRV and baroreflex sensitivity in obese individuals and in older age groups of normal-weight participants (Figure 3). These alterations are signs of arterial stiffening and a loss of cardiovagal control that can be observed in elderly individuals (Pinto, 2007). Advanced cardiovascular aging was suggested by an age gap over 5 years when compared to matched normal-weight controls.

The relationship between body mass index (BMI) and mortality is well documented (see review by Aune et al., 2016). While increased BMI raises mortality risk (Chen et al., 2019), a large population-based study has recently demonstrated that weight loss can prevent premature death in later life (Xie et al., 2020). Participants who reduced their BMIs below the obese range between early adulthood through midlife halved their mortality risk compared with those remaining obese, suggesting that the physiological effects of obesity may be reversible to some extent. Expressing cardiovascular impairment in terms of advanced age might help to convince those individuals at risk to adopt a healthier lifestyle (Cuende, 2016).

Limitations

The current study relies on physically and mentally healthy subjects who were recruited for resting physiological recordings under standardized conditions. However, the size of the sample is, therefore, rather small. Although we investigated over a thousand subjects, the number of data sets actually included in the analysis was reduced by quality control. Especially at older ages, a rather small amount of data was available. The recruitment of participants of an advanced age without being affected by cardiovascular, neurological, or psychiatric disorders is very complicated. Further cognitive impairment, sensory loss, and changes in mobility might introduce a selection bias (Young and Vitaliano, 2006).

Another limitation of our database is that there is no information on general health in order to account for it in our analysis, such as metabolic markers, smoking or drinking habits, or mental health. Because we also lack longitudinal data, we are unable to evaluate how autonomic status continues to change. Aging of the cardiovascular system is, most probably, not an entirely linear process. Intercurrent life events might moderate the rate of age-related changes. An intriguing line of further research is to assess the potential of the estimated age to predict cardiovascular risk.

Conclusion

In this study, we estimated age based on autonomic cardiovascular indices with high accuracy in healthy controls. The Gaussian process regression model led to the best concordance of estimated and calendar age. Using this framework, it seems possible to quantify deviations from healthy autonomic aging. In this study, cardiovascular changes in obese but otherwise healthy individuals led to an advanced age of more than 5 years compared with normal-weight controls. In future studies, the clinical value of the gap between the individual calendar and the estimated autonomic age to indicate diseases of the circulatory system or its potential to predict cardiovascular risk needs to be explored.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: https://physionet.org/content/autonomic-aging-card; https://doi.org/10.13026/2hsy-t491

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the University Hospital of Jena. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AS performed data analysis and wrote the manuscript. CG supervised analysis and interpretation of data. PCS advised the interpretation of results and critically revised the manuscript. RS, SF, and CS supervised analysis and critically revised the manuscript. K-JB conceived the study and prepared and critically

revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the German Research Foundation (DFG, BA 3848/9-1 and SCHU 3432/2-1) and the Interdisciplinary Centre for Clinical Research Jena (IZKF, MSP05-2019).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.899249/full#supplementary-material

References

Acevedo, M. A., Corrada-Bravo, C. J., Corrada-Bravo, H., Villanueva-Rivera, L. J., and Aide, T. M. (2009). Automated classification of bird and amphibian calls using machine learning: a comparison of methods. *Eco. Inform.* 4, 206–214. doi: 10.1016/j.ecoinf.2009.06.005

Al Zoubi, O., Al Zoubi, O., Ki Wong, C., Kuplicki, R. T., Yeh, H. W., Mayeli, A., et al. (2018). Predicting age from brain EEG signals-a machine learning approach. *Front. Aging Neurosci.* 10, 1–12. doi: 10.3389/fnagi.2018.00184

AlGhatrif, M., Strait, J. B., Morrell, C. H., Canepa, M., Wright, J., Elango, P., et al. (2013). Longitudinal trajectories of arterial stiffness and the role of blood pressure: the Baltimore longitudinal study of aging. *Hypertension* 62, 934–941. doi: 10.1161/HYPERTENSIONAHA.113.01445

Attia, Z. I., Friedman, P., Noseworthy, P., Lopez-Jimenez, F., Ladewing, G., Satam, G., et al. (2019). Age and sex estimation using artificial intelligence from standard 12-Lead ECGs. *Circ. Arrhythm. Electrophysiol.* 12, 1–11. doi: 10.1161/CIRCEP.119.007284

Aune, D., Sen, A., Prasad, M., Norat, T., Janszky, I., Tonstad, S., et al. (2016). BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *BMJ* 353:i2156. doi: 10.1136/bmj.i2156

Bauer, A., Kantelhardt, J. W., Barthel, P., Schneider, R., Mäkikallio, T., Ulm, K., et al. (2006). Deceleration capacity of heart rate as a predictor of mortality after myocardial infarction: cohort study. *Lancet* 367, 1674–1681. doi: 10.1016/S0140-6736(06)68735-7

Baumert, M., Baier, V., Haueisen, J., Wessel, N., Meyerfeldt, U., Schirdewan, A., et al. (2004). Forecasting of life threatening arrhythmias using the compression entropy of heart rate. *Methods Inf. Med.* 43, 202–206. doi: 10.1055/s-0038-1633859

Bennett, K. P., and Campbell, C. (2000). Support vector machines: hype or hallelujah? *ACM SIGKDD Explor. Newsl.* 2, 1–13. doi: 10.1145/380995.380999

Berger, R. D. (2003). QT Variability. *J. Electrocardiol.* 36, 83–87. doi: 10.1016/j.jelectrocard.2003.09.019

Boettger, M. K., Schulz, S., Berger, S., Tancer, M., Yeragani, V. K., Voss, A., et al. (2010). Influence of age on linear and nonlinear measures of autonomic cardiovascular modulation. *Ann. Noninvasive Electrocardiol.* 15, 165–174. doi: 10.1111/j.1542-474X.2010.00358.x

Chen, M., Hai, Y., Hwang, K., Wang, L., and Wang, L. (2017). Disease prediction by machine learning from healthcare communities. *IEEE Access* 5, 8869–8879. doi: 10.1109/ACCESS.2017.2694446

Chen, C., Ye, Y., Zhang, Y., Pan, X. F., and Pan, A. (2019). Weight change across adulthood in relation to all cause and cause specific mortality: prospective cohort study. *BMJ* 367:l5584. doi: 10.1136/bmj.l5584

Cole, J. H., and Franke, K. (2017). Predicting age using neuroimaging: innovative brain ageing biomarkers. *Trends Neurosci.* 40, 681–690. doi: 10.1016/j.tins.2017.10.001

Colosimo, A. (1997). Estimating a cardiac age by means of heart rate variability. *Am. J. Physiol. Heart Circ. Physiol.* 273, H1841–H1847. doi: 10.1152/ajpheart.1997.273.4.H1841

Cuende, J. I. (2016). Vascular Age Versus Cardiovascular Risk: Clarifying Concepts. *Rev. Esp. Cardiol.* 69, 243–246. doi: 10.1016/j.recesp.2015.10.022

Dafflon, J., Pinaya, W. H. L., Turkheimer, F., Cole, J. H., Leech, R., Harris, M. A., et al. (2020). An automated machine learning approach to predict brain age from cortical anatomical measures. *Hum. Brain Mapp.* 41, 3555–3566. doi: 10.1002/hbm.25028

Dai, X., Hummel, S. L., Salazar, J. B., Taffet, G. E., Zieman, S., and Schwartz, J. B. (2015). Cardiovascular physiology in the older adults. *J. Geriatr. Cardiol.* 12, 196–201. doi: 10.11909/j.issn.1671-5411.2015.03.015

Fischer, C., Glos, M., Penzel, T., and Fietze, I. (2017). Extended algorithmfor real-time pulse waveform segmentation and artifact detection in photoplethysmograms. *Somnologie* 21, 110–120. doi: 10.1007/s11818-017-0115-7

Floras, J. S. (2013). Blood pressure variability: a novel and important risk factor. *Can. J. Cardiol.* 29, 557–563. doi: 10.1016/j.cjca.2013.02.012

Franke, K., and Gaser, C. (2019). Years of brainage as a neuroimaging biomarker of brain aging: what insights have we gained? *Front. Neurol.* 10:789. doi: 10.3389/fneur.2019.00789

Fukusaki, C., Kawakubo, K., and Yamamoto, Y. (2000). Assessment of the primary effect of aging on heart rate variability in humans. *Clin. Auton. Res.* 10, 123–130. doi: 10.1007/BF02278016

Gialluisi, A., Di Castelnuovo, A., Donati, M. B., de Gaetano, G., and Iacoviello, L. (2019). Machine learning approaches for the estimation of biological aging: the road ahead for population studies. *Front. Med.* 6, 1–7. doi: 10.3389/fmed.2019.00146

Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101, e215–e220. doi: 10.1161/01.CIR.101.23.e215

Hodges, M., Salerno, D., and Erlien, D. (1983). Bazett's QT correction reviewed: evidence that a linear QT correction for heart rate is better. *J. Am. Coll. Cardiol.* 1, 577–739. doi: 10.1016/S0735-1097(83)80095-3

Jensen, M. T., Suadicani, P., Hein, H. O., and Gyntelberg, F. (2013). Elevated resting heart rate, physical fitness and all-cause mortality: a 16-year follow-up in the Copenhagen male study. *Heart* 99, 882–887. doi: 10.1136/heartjnl-2012-303375

Jiang, H., Lu, N., Chen, K., Yao, L., Li, K., Zhang, J., et al. (2020). Predicting brain age of healthy adults based on structural MRI Parcellation using convolutional neural networks. *Front. Neurol.* 10:1346. doi: 10.3389/fneur.2019.01346

Kakadiaris, I. A., Vrigkas, M., Yen, A. A., Kuznetsova, T., Budoff, M., and Naghavi, M. (2018). Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *J. Am. Heart Assoc.* 7:e009476. doi: 10.1161/JAHA.118.009476

Koch, M. (2018). Artificial intelligence is becoming natural. *Cells* 173, 531–533. doi: 10.1016/j.cell.2018.04.007

Laitinen, T., Niskanen, L., Geelen, G., Länsimies, E., and Hartikainen, J. (2004). Age dependency of cardiovascular autonomic responses to head-up tilt in healthy subjects. *J. Appl. Physiol.* 96, 2333–2340. doi: 10.1152/japplphysiol.00444.2003

Malberg, H., Wessel, N., Schirdewan, A., Osterziel, K. J., and Voss, A. (1999). Dual sequence method for analysis of spontaneous baroreceptor reflex sensitivity in patients with dilated cardiomyopathy. *Z. Kardiol.* 88, 331–337. doi: 10.1007/s003920050294

Malik, M., Bigger, J., Camm, A., and Kleiger, R. (1996). Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task force of the European Society of Cardiology and the north American Society of Pacing and Electrophysiology. *Eur. Heart J.* 17, 354–381. doi: 10.1093/oxfordjournals.eurheartj.a014868

Mendis, S., Puska, P., and Norrving, B. (2011). *Global atlas on cardiovascular disease prevention and control* Geneva, Switzerland: World Health Organization, 2–14.

Nanchen, D., Stott, D. J., Gussekloo, J., Mooijaart, S. P., Westendorp, R. G. J., Jukema, J. W., et al. (2013). Resting heart rate and incident heart failure and cardiovascular mortality in older adults: role of inflammation and endothelial dysfunction: the PROSPER study. *Eur. J. Heart Fail.* 15, 581–588. doi: 10.1093/eurjhf/hfs195

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830. doi: 10.48550/arXiv.1201.0490

Pinto, E. (2007). Blood pressure and ageing. *Postgrad. Med. J.* 83, 109–114. doi: 10.1136/pgmj.2006.048371

Reardon, M., and Malik, M. (1996). Changes in heart rate variability with age. *PACE Pacing Clin. Electrophysiol.* 19, 1863–1866. doi: 10.1111/j.1540-8159.1996.tb03241.x

Ren, X., Mi, Z., and Georgopoulos, P. G. (2020). Comparison of machine learning and land use regression for fine scale spatiotemporal estimation of ambient air pollution: modeling ozone concentrations across the contiguous United States. *Environ. Int.* 142, 105827–105813. doi: 10.1016/j.envint.2020.105827

Renyi, A. (1961). *On measures of entropy and information. In fourth Berkeley symposium on mathematical statistics and probability*. Berkeley, CA, USA: University of California Press.

Richman, J. S., and Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy maturity in premature infants physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* 278, H2039–H2049. doi: 10.1152/ajpheart.2000.278.6.H2039

Schmalenberger, K. M., Eisenlohr-Moul, T., Jarczok, M., Eckstein, M., Schneider, E., Brenner, I., et al. (2020). Menstrual cycle changes in vagally-mediated heart rate variability are associated with progesterone: evidence from two within-person studies. *J. Clin. Med.* 9, 1–20. doi: 10.3390/jcm9030617

Schölkopf, B., and Smola, A. (2002). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. Cambridge, MA: MIT Press.

Schulz, E., Speekenbrink, M., and Krause, A. (2018). A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *J. Math. Psychol.* 85, 1–16. doi: 10.1016/j.jmp.2018.03.001

Schumann, A., and Bär, K.-J. (2022). Autonomic aging – a dataset to quantify changes of cardiovascular autonomic function during healthy aging. *Sci. Data* 9, 1–5. doi: 10.1038/s41597-022-01202-y

Sevakula, R. K., AU-Yeung, M., Singh, J., Heist, K., Isselbacher, E. M., and Armoundas, A. (2020). State-of-the-art machine learning techniques aiming to improve patient outcomes pertaining to the cardiovascular system. *J. Am. Heart Assoc.* 9, 1–15. doi: 10.1161/JAHA.119.013924

Starc, V., Lebon, M., Sinigoj, P., and Vrhovec, M. (2012). Can functional cardiac age be predicted from the ECG in a normal healthy population? *Comput. Cardiol.* 2010, 101–104.

Strodthoff, N., Wagner, P., Schaeffter, T., and Samek, W. (2021). Deep learning for ECG analysis: benchmarks and insights from PTB-XL. *IEEE J. Biomed. Heal. Informatics* 25, 1519–1528. doi: 10.1109/JBHI.2020.3022989

Suri, J. S., Bhagawati, M., Paul, S., Protogeron, A., Sfikakis, P. P., Kitas, G. D., et al. (2022). Understanding the bias in machine learning systems for cardiovascular disease risk assessment: the first of its kind review. *Comput. Biol. Med.* 142:105204. doi: 10.1016/j.compbiomed.2021.105204

Tipping, M. E. (2000). The relevance vector machine. in *Advances in neural information processing systems 12* (ed. S. A. Solla and T. K. Leen, and M, K.-R.) 652–658. Cambridge, MA, USA: MIT Press.

Tsuji, H., Larson, M. G., Venditti, F. J., Manders, E. S., Evans, J. C., Feldman, C. L., et al. (1996). Impact of reduced heart rate variability on risk for cardiac events. *Circulation* 94, 2850–2855. doi: 10.1161/01.CIR.94.11.2850

Vapnik, V. (1995). *The nature of statistical learning theory* New York, NY, USA: Springer.

Voss, A., Heitmann, A., Schroeder, R., Peters, A., and Perz, S. (2012). Short-term heart rate variability - age dependence in healthy subjects. *Physiol. Meas.* 33, 1289–1311. doi: 10.1088/0967-3334/33/8/1289

Voss, A., Schroeder, R., Heitmann, A., Peters, A., and Perz, S. (2015). Short-term heart rate variability - influence of gender and age in healthy subjects. *PLoS One* 10, 1–33. doi: 10.1371/journal.pone.0118308

Wessel, N., Voss, A., Malberg, H., Ziehmann, C., Voss, H. U., Schirdewan, A., et al. (2000). Nonlinear analysis of complex phenomena in cardiological data. *Herzschrittmacherther. Elektrophysiol.* 11, 159–173. doi: 10.1007/s003990070035

Xie, W., Lundberg, D. J., Collins, J. M., Johnston, S. S., Waggoner, J. R., Hsiao, C. W., et al. (2020). Association of Weight Loss between Early Adulthood and Midlife with all-Cause Mortality Risk in the US. *JAMA Netw. Open* 3, 1–11. doi: 10.1001/jamanetworkopen.2020.13448

Young, H., and Vitaliano, P. (2006). "Methods in Health Psychology" in *Handbook of the psychology of aging*, Carolyn M. Aldwin, Crystal L. Park, and Avron Spiro. New York, NY, USA: Guilford Press, 54–74.

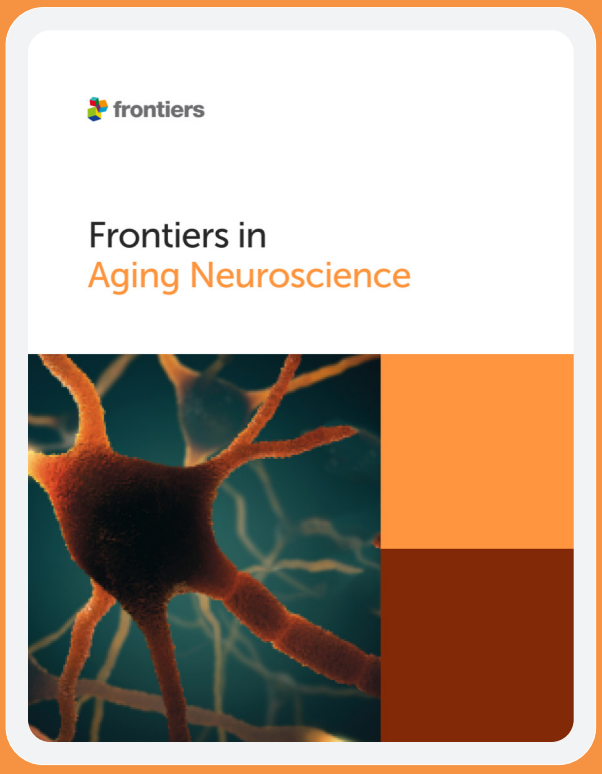
Frontiers in
Aging Neuroscience

Explores the mechanisms of central nervous system aging and age-related neural disease

The third most-cited journal in the field of geriatrics and gerontology, with a focus on understanding the mechanistic processes associated with central nervous system aging.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact