



Original Article

Structural Alterations and Cognitive Impairment in Late-Onset Depression: A Reverse Correlation Analysis

Yujie Wu^{1,2}, Naikeng Mai^{3,*}¹School of Education Science, Guangdong Polytechnic Normal University, 510665 Guangzhou, Guangdong, China²Department of Psychosomatic Medicine, Guangdong Provincial People's Hospital, 510080 Guangzhou, Guangdong, China³Department of Neurology, The Affiliated Brain Hospital of Guangzhou Medical University (Guangzhou Huiai Hospital), 510370 Guangzhou, Guangdong, China*Correspondence: mnkmaksin@163.com (Naikeng Mai)

Academic Editor: Francesco Bartoli

Submitted: 6 July 2025 Revised: 25 October 2025 Accepted: 10 November 2025 Published: 3 February 2026

Abstract

Background: Late-onset depression (LOD), particularly when accompanied by cognitive impairment, represents a significant risk factor for dementia. Prevailing perspectives emphasize that cognitive impairment arises from interactions among multiple brain regions. However, current approaches to identifying brain network patterns associated with cognitive impairment largely rely on group-level analyses with multiple-comparison corrections, which may obscure complex and interconnected relationships between brain regions. Our previous research demonstrated that alterations in brain network properties in patients with LOD are closely associated with cognitive function. We therefore hypothesised that aberrant interactions among multiple brain regions in LOD lead to changes in network properties and subsequent cognitive dysfunction. **Methods:** This study aimed to investigate the interregional brain interactions underlying cognitive impairment in LOD by leveraging the robust interpretability of neural network models. Specifically, we sought to: (1) develop a neural network model of LOD-related cognitive impairment based on brain network properties; and (2) apply a reverse correlation approach to identify connectivity features associated with cognitive impairment in LOD. **Results:** No statistically significant differences were observed in the structural network properties when comparing the LOD and control participant groups across various thresholds. Using a neural network-based reverse correlation method, the most prominent differences were identified in the inferior, middle, and anterior regions of the left temporal pole when comparing patients with LOD with and without mild cognitive impairment (MCI). **Conclusion:** Alterations in the internal structure of the temporal lobe may represent potential anatomical biomarkers for the early prediction of Alzheimer's disease, providing novel insights into the pathophysiological mechanisms underlying LOD-related MCI. The research framework proposed in this study effectively addresses the challenge of detecting subtle intergroup anatomical differences in studies with limited sample sizes. Moreover, the reverse correlation approach is not restricted to multilayer neural networks; as machine learning models become increasingly powerful and accessible, this method offers a practical and interpretable alternative for exploratory neuroimaging research.

Keywords: depression; cognitive impairment; machine learning; neural networks; magnetic resonance imaging

Main Points

1. Patients with late-onset depression and mild cognitive impairment exhibit specific structural alterations in brain regions.
2. The neural network and reverse correlation-based approach effectively identifies subtle intergroup differences.
3. Temporal lobe structural alterations may reflect an early stage in the transition from late-onset depression to cognitive impairment.

1. Introduction

With an ageing population, the healthcare burden related to late-life conditions continues to rise. Late-life depression (LLD), defined as depression occurring in adults aged 60 and above, is one of the most common mental disorders among older adults. The risk of depression increases

with age in this group. LLD is a known risk factor for Alzheimer's disease and vascular dementia [1], and when coexisting with mild cognitive impairment (MCI), it significantly predicts dementia [2–4]. LLD can be divided into early-onset depression, which recurs or persists into old age, and late-onset depression (LOD), characterised by a first depressive episode occurring at 60 or older. early-onset depression may be associated with genetic predisposition, a family history of mental illness, and stress-related factors [5,6]. Our recent study identified distinct differences between early-onset depression and LOD in dynamic network modularity and rich-club node distribution, assessed via resting-state functional magnetic resonance imaging [7]. These differences correlated with cognitive function in only one subtype, suggesting distinct mechanisms of cognitive impairment in early-onset depression and LOD. Given that LOD poses a higher risk for cognitive decline than early-onset depression, this study focused on analysing the



structural correlates of cognitive impairment specifically in LOD.

Graph theory enables the quantitative analysis of brain network topology. The key metrics in this respect include: global efficiency; nodal strength; clustering coefficient; shortest path length; and local efficiency. These metrics capture how efficiently information is integrated both globally and locally, and relate closely to cognitive abilities and clinical symptoms. Rich-club properties and modularity further describe the brain's balance between integration and segregation.

Our previous work demonstrated synergistic effects of LLD and MCI on brain ageing. Structural covariance networks revealed that grey matter alterations in LLD-MCI resemble those observed in Alzheimer's disease [8]. Memory impairment in LLD correlates with white matter rich-club reorganisation, where the connectivity strength of rich-club and local networks, shortest path length, global efficiency, and local efficiency, is associated with cognitive function [9], indicating both grey and white matter involvement in LLD-related cognitive decline. Although functional magnetic resonance imaging (fMRI) study in LLD-MCI have reported similar network alterations, no cognitive correlations were identified [10]. Our recent dynamic network analysis showed that modularity correlates with memory in early-onset depression but not in LOD, with similar patterns observed in rich-club analysis (see Research Foundation). This suggests distinct structural substrates for cognitive impairment in early-onset depression versus LOD. However, the specific brain regions driving changes in rich-club organization and modularity remain unidentified. Pinpointing the key nodes influencing rich-club organization is crucial for understanding the neuropathology of LLD and for developing targeted neuromodulation therapies, such as Transcranial Magnetic Stimulation/transcranial Direct Current Stimulation.

Advances in methodology have broadened the applications of machine learning [11]. Clinically, machine learning supports diagnosis, prognosis prediction, and the assessment of treatment response [12–14]. It also enables the quantification of disease-induced neural damage [15]. Our preliminary studies applied machine learning to evaluate brain ageing, confirming synergistic LLD-MCI effects on cognitive decline [8]. Although MRI-derived network properties partially reflect cognitive function in LLD [9], conventional comparisons fail to pinpoint the structural signatures of impairment. In this study, we employed neural networks to investigate the structural and functional features of cognitive impairment in LOD.

Owing to the strong interpretability of back propagation neural networks, a classification model can be established. However, unlike regression models, the presence of hidden layers in back propagation networks makes it challenging to interpret connection weights and classifier thresholds in real-world terms. Nevertheless, their excel-

lent generalisation performance makes back propagation networks a powerful tool for gaining deeper insights into underlying principles of complex systems. To explore their potential as feature-extraction tools in brain networks, we applied the reverse correlation method, using classification results to infer internal representations. Reverse correlation is a data-driven technique for visualising mental representations. In this approach, participants are shown a series of stimuli created by adding random noise patterns to a base image, typically a face with a neutral expression. These variations were generated by superimposing random noise patterns onto the base image [16].

In this study, the neural network model served as the “participant” classifying healthy human brain networks that had been perturbed by the addition of random noise. By integrating these noise patterns, we extracted brain network features associated with cognitive impairment.

Given that interregional brain connectivity underpins higher-order neural functions, this study leveraged the predictive power of neural networks to identify dysregulated connectivity in LOD-related cognitive impairment. Using reverse correlation, we extracted discriminative features from the model to uncover structural network signatures of LOD pathology.

2. Methods

This cross-sectional study was conducted at the Brain Hospital affiliated with Guangzhou Medical University, China. This study was approved by the ethics committee of the Guangdong Provincial People's Hospital and written informed consent was obtained from each participant. The inclusion criteria for patients with LOD were as follows: (1) age ≥ 60 years; (2) diagnosis of major depressive disorder according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5); (3) confirmation of major depressive disorder by psychiatrists at the Brain Hospital affiliated with Guangzhou Medical University; (4) the ability to complete neuropsychological assessments; (5) first depressive episode occurring at or after the age of 60. The exclusion criteria included: (1) the inability to complete neuropsychological testing; (2) a history of severe psychiatric disorders (e.g., bipolar disorder and schizophrenia) excluding depression; (3) a family history of severe psychiatric disorders (e.g., bipolar disorder and schizophrenia) excluding depression; (4) primary neurological disorders (e.g., stroke and brain tumours); (5) medical conditions potentially affecting mood (e.g., hypothyroidism, syphilis, and anaemia); (6) contraindications for MRI (e.g., non-cooperation or intracorporeal metallic implants); (7) Hamilton rating scale for depression (HAM-D) > 7 ; (8) Clinical Dementia Rating (CDR) score > 0.5 ; (9) Hachinski Ischaemic Scale score > 4 .

Patients with LOD were further stratified into the following groupings:

MCI group: diagnosed using Petersen's 1999 criteria [17]: (1) self- or informant-reported memory complaints; (2) objective cognitive impairment ≥ 1.5 standard deviation units below age- and education-matched norms; (3) CDR = 0.5 and Global Deterioration Scale stages 2 and 3; (4) preserved activities of daily living; and (5) absence of dementia.

Cognitively normal group (NON-MCI): patients with LOD not meeting the MCI criteria.

Healthy controls (HCs) met the following criteria: (1) age >60 years with no history of depression; (2) exclusion criteria similar to the LOD group; and (3) CDR ≤ 0.5 (to exclude dementia-depression comorbidity).

The final study cohort was comprised of the following participants: (1) 42 patients with LOD and 90 HCs from prior studies who met all criteria; (2) one patient with LOD who was excluded owing to incomplete diffusion tensor imaging (DTI) scanning.

All participants underwent clinical assessments, neuropsychological testing, and neuroimaging within 1 week of enrolment.

2.1 Clinical Assessments

Diagnostic reviews were conducted by trained clinicians holding intermediate-to-senior professional titles, all certified in the Structured Clinical Interview for DSM-5 Axis I Disorders. Detailed demographic and clinical data were collected from all participants, and comprehensive physical examinations were performed to assess exclusion criteria. Screening for enrollment utilized the CDR, Mini-Mental State Examination (MMSE), and HAMD. The medication history of LOD are presented in **Supplementary Table 1**.

2.2 Neuroimaging Data Acquisition

Cranial MRI scans were performed using a 3.0-Tesla MRI system (Siemens Magnetom PRISMA, Siemens Healthcare, Erlangen, Germany) at the Guangzhou Medical University Affiliated Brain Hospital, China. Prior to the formal scanning protocol, T2-weighted imaging was conducted to exclude cerebral infarction, haemorrhage, tumours, malformations, and extensive white matter lesions.

Prior to resting-state functional magnetic resonance imaging scanning, participants were instructed to remain relaxed with their eyes closed while refraining from falling asleep. Resting-state functional magnetic resonance imaging images were acquired using a gradient-echo echo-planar imaging sequence with the following parameters: TR = 2000 ms, TE = 30 ms, FA = 90° , matrix = 64×64 , slice thickness = 4 mm, slice gap = 0.6 mm, 33 interleaved axial slices, field of view (FOV) = 220×220 mm², and 240 time points.

DTI parameters were as follows: 32 directions, b -value = 1000 s/mm², TR = 10,015 ms, TE = 92 ms, flip angle = 90° , imaging matrix = 128×128 , FOV = $256 \times$

256 mm², 75 contiguous slices, and voxel dimensions = $2 \times 2 \times 2$ mm³.

High-resolution T1-weighted images were acquired using a three-dimensional (3D) spoiled gradient-echo sequence with the following parameters: TR = 8.2 ms, TE = 3.8 ms, slice thickness = 1 mm, FOV = 256×256 mm², and matrix = $256 \times 256 \times 188$.

2.3 Brain Network Construction and Feature Computation Data Preprocessing

DTI preprocessing was performed using the FMRIB Software Library (FSL) Diffusion Toolbox and included the following steps: Eddy current correction: compensated for magnetic field distortions and head motion. (1) b_0 extraction and skull-stripping: intensity threshold = 0.2. (2) Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques (BedpostX): generated whole-brain voxel-wise fibre orientation distributions. (3) T1 skull-stripping: intensity threshold = 0.3.

2.4 Network Construction

Node Definition: Using the Desikan–Killiany atlas, the following steps were performed: (1) T1 structural images were non-linearly registered to the MNI152 standard brain template (based on 3D MRI scans of 152 healthy individuals), producing an inverse warp file. (2) The Desikan–Killiany atlas was then transformed from T1 space to b_0 space using the corresponding transformation matrix. (3) This resulted in 82 regional masks registered to each subject's DTI space. Each regional mask served as a seed, while the remaining 81 masks were treated as terminal masks. Edge definition: probabilistic fibre tracking (FSL 5.09) was employed to define network edges. For each voxel within a seed mask, 5000 streamlines were generated. Streamline propagation followed the orientation distributions estimated by BedpostX, advancing in 0.5-mm steps toward the 81 target masks. Tracking stopped upon reaching a target mask to prevent infinite loops caused by circular fibres.

The connection probability between the seed mask and each target mask was calculated and used as the connection weight between nodes. After tracking was completed for all 82 seed masks, an 82×82 connectivity matrix representing individual white matter structural connectivity was generated. As probabilistic tracking does not provide directionality, the connection weights between nodes i and j (weight_{ij} and weight_{ji}) were used to produce an undirected connection weight.

2.5 Network Calculation

Network properties were computed across multiple connection density thresholds using custom scripts. The area under the curve (AUC) for each property across thresholds represented the overall network characteristic of the participant. The following parameters were cal-

culated in this study: clustering coefficient (aCp), global efficiency (aEg), local efficiency ($aEloc$), characteristic path length (aLp), modularity (aQ), rich-club connection strength ($Rich_s$), local connection strength ($Local_s$), and feeder connection strength ($Feeder_s$). Equations and detailed interpretations have been provided in our previous study [9].

2.6 Construction of a Neural Network Model for Cognitive Impairment Based on Brain Network Features

Features including aQ , $Local_s$, $Rich_s$, aEg , $aEloc$, and aLp were computed for both early-onset depression and LOD groups using the Brain Connectivity Toolbox (<https://sites.google.com/site/bctnet/home>) and custom Matrix Laboratory (MATLAB) (2024b, <https://ww2.mathworks.cn/products/matlab.html>) scripts. These features were then used to train a cognitive prediction model. Labels (cognitively normal vs. impaired) were assigned according to Petersen's 1999 MCI diagnostic criteria.

The neural network consisted of three hidden layers with six units each. Data were randomly split into training, testing, and validation sets in an 8:1:1 ratio. Implemented using the Neural Network Toolbox (3.0, The MathWorks, Inc., Natick, MA, USA) of MATLAB, the model employed: (1) training algorithm: scaled conjugate gradient; (2) performance metric: cross-entropy; and (3) given the limited feature set, no regularisation was applied to hidden layers.

The 3D distribution of real versus predicted data and the receiver operating characteristic curve for the LOD cognitive impairment classification model are presented in **Supplementary Figs. 1,2**.

2.6.1 Standard Network Construction

The structural connectivity matrices from 90 cognitively normal older participants were used to construct an edge-weight distribution curve for the HC group. After removing the top and bottom 2.5% of values, the mean of the remaining edge weights was calculated to establish a reference matrix for healthy ageing brain networks.

2.6.2 Randomisation Procedure

A difference matrix was generated by comparing the data of the LOD with that of the HC group. This difference matrix was randomised using the procedure depicted in Fig. 1 to produce randomised network matrices. Modular and rich-club partitions were then performed, and the normalised mutual information (NMI) for these partitions was calculated for both LOD and HC configurations derived from the randomised matrices. Randomised matrices satisfying the criterion $NMI_{LOD} < NMI_{HC}$ were selected for further analysis. The equation used to compute NMI is given below.

$$NMI(A1, A2) = \frac{-2 \sum_{i=1}^{C_{A1}} \sum_{j=1}^{C_{A2}} N_{ij} \log \left(\frac{N_{ij} N}{N_i N_j} \right)}{\sum_{i=1}^{C_{A1}} N_i \log \left(\frac{N_i}{N} \right) + \sum_{j=1}^{C_{A2}} N_j \log \left(\frac{N_j}{N} \right)}.$$

In the normalized mutual information (NMI) calculation, A1 and A2 represent the modular and rich-club partitions from the randomised networks and the LOD (or HC) network, respectively. CA1 and CA2 indicate the number of partitions in A1 and A2. N_{ij} is the number of nodes shared between a partition in A1 and one in A2. The total number of nodes, N, is 82 in this study. N_i and N_j refer to the number of nodes within a partition of A1 and A2, respectively. A higher NMI value reflects greater similarity between the modular and rich-club partitions of the randomised and LOD (or HC) networks.

2.6.3 Reverse Correlation for Cognitive Impairment Network Signatures

A total of 10,000 randomised networks were generated. (1) The NMI coefficient was calculated based on network properties and compared with the NMI coefficients of LOD and HC in the actual samples. Randomised networks resembling the network properties of LOD were retained. (2) A pre-trained neural network model was then used to classify these retained randomised networks, extracting those identified as MCI or NON-MCI. (3) Randomised difference matrices were derived from the networks classified in step 2. Aggregating these matrices revealed network connections associated with MCI or NON-MCI risk. To identify meaningful differences within the limited sample size, the top 5% most frequently occurring connections were selected as potential LOD network features related to MCI or NON-MCI.

2.7 Statistical Analysis

Differences in age, HAMD, MMSE, and CDR between patients with LOD and HCs were analysed using the independent-samples *t*-tests. Differences in sex distribution between groups were assessed using the Chi-square test. Group differences in structural network properties were evaluated using multivariate analysis of covariance, controlling for age as a covariate. For demographic data that did not pass the Levene's test, the Wilcoxon rank-sum test was employed. For group differences in structural network characteristics that violated normality assumptions, the ranks of covariates and dependent variables were first calculated. Linear regression was then used to adjust the ranks of the dependent variable based on the ranks of the independent variable, followed by the Kruskal–Wallis test to compare the adjusted residuals. The significance level was set at $\alpha = 0.05$, with $p < 0.05$ considered statistically significant.

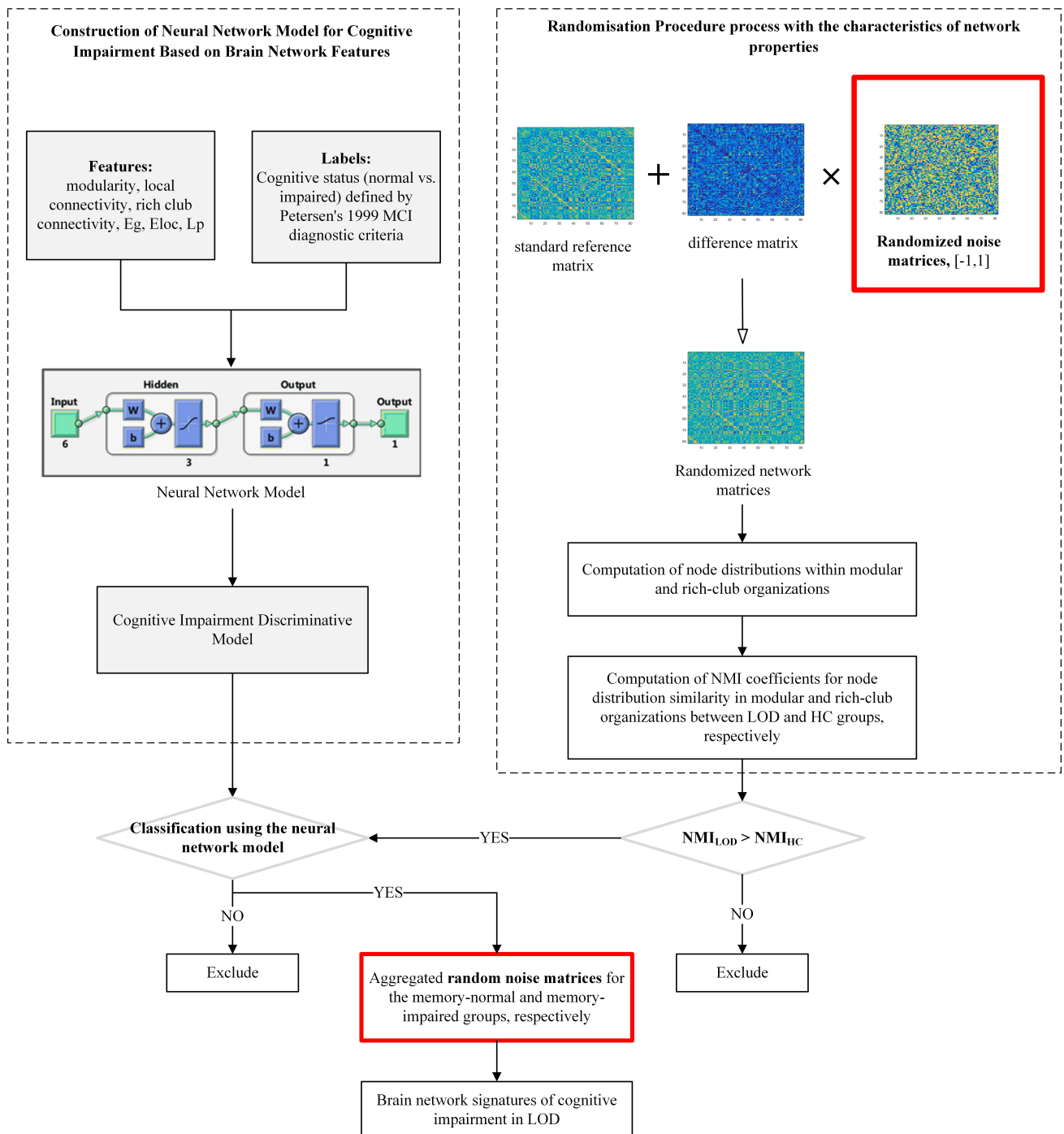


Fig. 1. Flowchart of the study on network characteristics of LOD with cognitive impairment using randomisation procedure and reverse correlation method. The red box represents the data source of the final result, which is a synthesis of random matrix data filtered by the discriminant model. LOD, Late-onset depression; MCI, mild cognitive impairment; NMI, normalised mutual information; HC, Healthy control.

3. Results

3.1 Demographic Characteristics

No statistically significant difference in age and sex distribution was observed between the LOD and HC groups ($p > 0.05$). However, significant differences were found in the HAMD, MMSE, and CDR scores between the two

groups ($p < 0.05$). No statistically significant difference in age, sex distribution, HAMD, MMSE, and CDR scores was observed between MCI and NON-MCI subgroups within the LOD ($p > 0.05$). The age of onset of LOD was 65.98 ± 5.99 . Detailed data are presented in Tables 1,2.

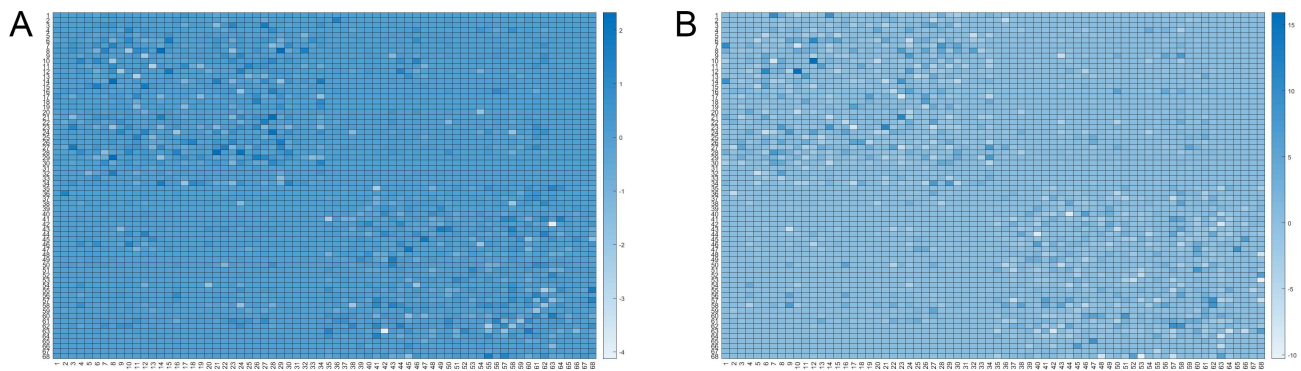


Fig. 2. Matrix visualisations depict the mean of difference node matrices derived from randomisation cycles ($n = 10,000$). For LOD-MCI, the mean of 2645 difference matrices classified as having cognitive impairment was used to generate visualisation (A). Similarly, for LOD non-MCI, the mean of 118 cognitively impaired-classified difference matrices was used to produce visualization (B). The color represents the frequency, ranging from white to dark blue, where a darker shade indicates a higher occurrence rate of the connection in the selected random matrices. Detailed node region information is provided in **Supplementary Material 1**.

Table 1. Demographic characteristics of LOD vs. HC groups.

	Age (years)	Sex (male/female)	HAMD	MMSE	CDR
LOD ($n = 41$)	68.09 ± 6.93	11/30	1.46 ± 2.15	27.26 ± 1.69	0.18 ± 0.24
HCs ($n = 90$)	69.93 ± 6.40	32/58	8.76 ± 7.73	24.34 ± 3.68	0.40 ± 0.20
$t/z/\chi^2$	$t = -1.441$	$\chi^2 = 0.972$	$z = -6.416$	$z = -4.646$	$z = -4.634$
df	129	1	129	129	129
p	0.152	0.324	0.001	0.001	0.001

HAMD, Hamilton rating scale for depression; MMSE, Mini-Mental State Examination; CDR, Clinical Dementia Rating.

Table 2. Demographic characteristics of MCI vs. NON-MCI subgroups within the LOD.

	Age (years)	Sex (male/female)	HAMD	MMSE	CDR
MCI ($n = 21$)	70.62 ± 6.92	5/16	10.29 ± 7.85	23.33 ± 3.61	0.45 ± 0.15
NON-MCI ($n = 20$)	69.20 ± 5.90	6/14	7.15 ± 7.46	25.40 ± 3.53	0.35 ± 0.24
$t/z/\chi^2$	$t = 0.705$	$\chi^2 = 0.200$	$t = 1.309$	$t = -1.852$	$z = -1.633$
df	39	1	39	39	39
p	0.485	0.655	0.198	0.072	0.102

NON-MCI, cognitively normal group.

3.2 Structural Network Metrics in LOD

No statistically significant differences were found in the structural network properties between the LOD and HC groups. Detailed data are presented in Table 3.

When comparing network features between MCI and NON-MCI subgroups within the LOD group, no statistically significant differences were detected in the AUC values of structural network properties across thresholds. Detailed data are presented in Table 4.

3.3 Node-Level Features in LOD Extracted via Reverse Correlation

For the LOD-MCI subgroup, 10,000 randomisation cycles produced a difference node matrix, resulting in 2645 significant difference matrices. The mean of these was used to produce the matrix visualisation shown in Fig. 2A. Simi-

larly, for the LOD-NON-MCI subgroup, 10,000 randomisation cycles resulted in 118 significant difference matrices, and the mean of these was used to generate the matrix in Fig. 2B.

A comparison of significant random matrix counts revealed that the structural networks in the LOD-MCI group exhibited the most pronounced deviations from the HC group. The most substantial differences were localised in the left inferior, middle, and anterior temporal pole regions. Detailed illustrations are provided in **Supplementary Material 2**.

4. Discussion

This study established a standard brain network matrix for healthy ageing using DTI structural images from 90 HCs. By integrating randomised matrices with the differ-

Table 3. Structural network metrics in LOD.

Network properties	<i>aCp</i>		<i>aEg</i>		<i>aEloc</i>		<i>aLp</i>		<i>aQ</i>		<i>Rich_s</i>		<i>Local_s</i>		<i>Feeder_s</i>	
	LOD	HCS	LOD	HCS	LOD	HCS	LOD	HCS	LOD	HCS	LOD	HCS	LOD	HCS	LOD	HCS
<i>n</i>	41	90	41	90	41	90	41	90	41	90	41	90	41	90	41	90
<i>M</i> ± SD	0.110 ± 0.009	0.111 ± 0.007	0.104 ± 0.004	0.103 ± 0.009	0.144 ± 0.007	0.144 ± 0.009	1.057 ± 0.040	1.071 ± 0.108	0.146 ± 0.010	0.149 ± 0.014	0.626 ± 0.016	0.628 ± 0.021	0.473 ± 0.033	0.479 ± 0.035	0.569 ± 0.022	0.572 ± 0.025
<i>F</i>	0.252		0.322		0.114		0.815		1.222		0.208		0.738		0.304	
<i>p</i>	0.617		0.571		0.736		0.368		0.271		0.649		0.392		0.582	
FDR	0.736		0.736		0.736		0.736		0.736		0.736		0.736		0.736	

M ± SD, mean ± standard deviation; *aCp*, clustering coefficient; *aEg*, global efficiency; *aEloc*, local efficiency; *aLp*, characteristic path length; *aQ*, modularity; *Rich_s*, rich-club connection strength; *Local_s*, local connection strength; *Feeder_s*, feeder connection strength.

Table 4. Structural network metrics between MCI and NON-MCI subgroups within the LOD group.

Network properties	<i>aCp</i>		<i>aEg</i>		<i>aEloc</i>		<i>aLp</i>		<i>aQ</i>		<i>Rich_s</i>		<i>Local_s</i>		<i>Feeder_s</i>	
	MCI	NON-MCI	MCI	NON-MCI	MCI	NON-MCI	MCI	NON-MCI	MCI	NON-MCI	MCI	NON-MCI	MCI	NON-MCI	MCI	NON-MCI
<i>n</i>	21	20	21	20	21	20	21	20	21	20	21	20	21	20	21	20
<i>M</i> ± SD	0.111 ± 0.008	0.109 ± 0.009	0.104 ± 0.003	0.103 ± 0.005	0.143 ± 0.005	0.144 ± 0.008	1.054 ± 0.032	1.061 ± 0.047	0.145 ± 0.009	0.147 ± 0.012	0.626 ± 0.017	0.627 ± 0.016	0.474 ± 0.031	0.472 ± 0.035	0.569 ± 0.020	0.568 ± 0.023
<i>F</i>	0.411		0.202		0.110		0.204		0.201		0.018		0.034		0.044	
<i>p</i>	0.525		0.656		0.742		0.654		0.656		0.895		0.855		0.835	
FDR	0.895		0.895		0.895		0.895		0.895		0.895		0.895		0.895	

ence network between LOD and HC groups and combining them with the standard matrix, we generated random networks constrained within known LOD connectivity parameters. By employing feature comparison with established LOD networks and neural network filtering, we identified LOD-characterised random networks. An aggregation of these matrices revealed distinct network features for LOD-MCI and LOD-NON-MCI. The core innovation of the present study lies in applying reverse correlation to identify those network features that determine training efficiency in well-trained neural networks. This study proposed the application of the reverse correlation method as a tool for interpreting machine learning and identifying potential discrepancy trends in preliminary experiments with limited sample sizes. We aimed to detect divergence trends in sample collection at an early stage in the research, analyse experimental samples based on research objectives, and promptly adjust relevant conditions for sample collection.

No structural network differences were observed between the LOD-MCI and LOD-NON-MCI subgroups. However, our previous brain-age prediction studies had identified significant disparities, suggesting that conventional methods fail to capture abnormalities in clinically distinct cohorts. In contrast, the reverse correlation-based neural network approach successfully detected differential inter-nodal connections. Although statistical validation is still pending, these preliminary results highlight the need for larger-scale investigations.

Matrix significance comparisons revealed the greatest structural network deviations between LOD-MCI (2645/10,000 matrices) and HC groups, and these findings are consistent with the results of prior research. The left inferior, middle, and anterior temporal pole showed the most pronounced differences, suggesting that disrupted intra-temporal white matter connectivity may serve as a potential anatomical biomarker to distinguish HC from LOD-MCI. Although studies on LLD-related cognitive impairment have implicated the cingulate cortex, hippocampus, and prefrontal cortex, these regions lack disease specificity, as they also appear in conditions such as frontotemporal dementia [18] and Lewy body dementia [19]. The encodement of memory depends on coordinated hippocampal engagement [20], frontal lobe integration of entorhinal-hippocampal inputs [21], and medial temporal-hippocampal spatial processing [22]. Given that LOD-MCI is a known risk factor for Alzheimer's disease, and considering the limited sample size of this study, our findings suggest that structural changes within the dominant temporal lobe may reflect an early stage in the transition from LOD to Alzheimer's disease. These preliminary results may also provide valuable clues for advancing our understanding of the pathogenesis of LOD-MCI.

The structure–function relationships of the human brain manifest as complex networks that enable efficient segregation and integration of information. Neuroimaging-

derived connectivity matrices capture these architectures and can be quantified using graph theory, where nodes represent anatomical or functional regions of interest and edges denote structural or functional connections [23,24]. Network properties have been shown to correlate with cognition and symptomatology: Li *et al.* [25] linked LLD white matter connectivity to processing speed, while another study reported associations between functional network metrics and depression severity [10]. Our replication of Cp/Eg differences between LLD and HC groups supports established findings. The absence of significant LOD subgroup differences may be attributable to limited statistical power, whereas machine learning with reverse correlation successfully detected subtle intergroup variations.

Limitations and Future Directions

This study has some limitations. First, it employed the reverse correlation method to identify potential differences between groups. No corresponding statistical values were generated because this approach does not rely on hypothesis testing and instead uses reverse inference from predictive models to uncover possible points of divergence. The accuracy of the differences identified is therefore dependent on the generalisation capability of the model. Second, owing to the small sample size, overfitting was inevitable during neural network training, which reduced the generalisability of the final model. As a result, the findings may not be readily extrapolated to broader populations. We are conducting further in-depth research on this method and cross-validating it with other statistical approaches. In subsequent studies, we aim to establish a statistically meaningful minimum sample size for this method.

Nevertheless, the research framework proposed in this study effectively addresses the challenges in detecting subtle intergroup differences when sample sizes are limited. Based on its methodological principles, we hypothesise that as sample sizes increase and the predictive performance of neural network training improves, the conclusions derived may generalise more effectively to real-world scenarios than those obtained using traditional statistical methods, thereby underscoring the value of this approach and its potential for further investigation. Further, the framework is not confined to multilayer neural networks. Given the emergence of models that provide better machine learning-based training performance at lower computational costs, the reverse correlation method proposed here could serve as a simpler alternative for research exploration.

5. Conclusion

This study utilised multilayer neural networks combined with a reverse correlation approach, demonstrating that LOD-MCI may be linked to structural alterations in the temporal lobe, thereby providing valuable insights into its pathological mechanisms. Importantly, the proposed framework illuminates the decision-making logic of mul-

tilayer neural networks and effectively tackles the difficulty of detecting subtle inter-group differences when sample sizes are limited. Further, the reverse correlation method is not specific to neural networks and can be applied across other domains. Given the growing accessibility of high-performance and cost-efficient machine learning models, the reverse correlation technique introduced in this study presents a practical and accessible alternative for exploratory research in the field.

Availability of Data and Materials

The data and materials used in this study are available from the corresponding author upon reasonable request.

Author Contributions

NM designed and performed the research study. YW analyzed the data and drafted the manuscript. Both authors contributed to critical revision of the manuscript for important intellectual content. Both authors read and approved the final manuscript. Both authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Guangdong Provincial People's Hospital (ethics approval number: KY2023-576-03). All participants or their legal guardians provided signed informed consent.

Acknowledgment

Thanks to all the peer reviewers for their opinions and suggestions.

Funding

This research was funded by Guangzhou Science and Technology Program - Basic and Applied Basic Research Project (2023A04J0509), Guangzhou Science and Technology Program - Basic and Applied Basic Research Project (202202011280), and Heyuan Philosophy and Social Sciences Planning Project (HYSK25Z16). The funding agency had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.31083/AP44585>.

References

- [1] Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds CF, 3rd. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. *The British Journal of Psychiatry: the Journal of Mental Science*. 2013; 202: 329–335. <https://doi.org/10.1192/bjp.bp.112.118307>.
- [2] Artero S, Ancelin ML, Portet F, Dupuy A, Berr C, Dartigues JF, *et al.* Risk profiles for mild cognitive impairment and progression to dementia are gender specific. *Journal of Neurology, Neurosurgery, and Psychiatry*. 2008; 79: 979–984. <https://doi.org/10.1136/jnnp.2007.136903>.
- [3] Lu PH, Edland SD, Teng E, Tingus K, Petersen RC, Cummings JL, *et al.* Donepezil delays progression to AD in MCI subjects with depressive symptoms. *Neurology*. 2009; 72: 2115–2121. <https://doi.org/10.1212/WNL.0b013e3181aa52d3>.
- [4] Edwards ER, Spira AP, Barnes DE, Yaffe K. Neuropsychiatric symptoms in mild cognitive impairment: differences by subtype and progression to dementia. *International Journal of Geriatric Psychiatry*. 2009; 24: 716–722. <https://doi.org/10.1002/gps.2187>.
- [5] Thapar A, Eyre O, Patel V, Brent D. Depression in young people. *Lancet (London, England)*. 2022; 400: 617–631. [https://doi.org/10.1016/S0140-6736\(22\)01012-1](https://doi.org/10.1016/S0140-6736(22)01012-1).
- [6] Harald B, Gordon P. Meta-review of depressive subtyping models. *Journal of Affective Disorders*. 2012; 139: 126–140. <https://doi.org/10.1016/j.jad.2011.07.015>.
- [7] Mai N, Wu Y, Zhong X, Chen B, Zhang M, Peng Q, *et al.* Increasing variance of rich-club nodes distribution in early onset depression according to dynamic network. *Brain Imaging and Behavior*. 2024; 18: 662–674. <https://doi.org/10.1007/s11682-023-00848-5>.
- [8] Mai N, Wu Y, Zhong X, Chen B, Zhang M, Ning Y. Determining the effects of LLD and MCI on brain decline according to machine learning and a structural covariance network analysis. *Journal of Psychiatric Research*. 2020; 126: 43–54. <https://doi.org/10.1016/j.jpsychires.2020.04.011>.
- [9] Mai N, Zhong X, Chen B, Peng Q, Wu Z, Zhang W, *et al.* Weight Rich-Club Analysis in the White Matter Network of Late-Life Depression with Memory Deficits. *Frontiers in Aging Neuroscience*. 2017; 9: 279. <https://doi.org/10.3389/fnagi.2017.00279>.
- [10] Li W, Douglas Ward B, Liu X, Chen G, Jones JL, Antuono PG, *et al.* Disrupted small world topology and modular organization of functional networks in late-life depression with and without amnesic mild cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry*. 2015; 86: 1097–1105. <https://doi.org/10.1136/jnnp-2014-309180>.
- [11] Myszczyńska MA, Ojames PN, Lacoste AMB, Neil D, Saffari A, Mead R, *et al.* Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nature Reviews. Neurology*. 2020; 16: 440–456. <https://doi.org/10.1038/s41582-020-0377-8>.
- [12] Redlich R, Opel N, Grotegerd D, Dohm K, Zaremba D, Bürger C, *et al.* Prediction of Individual Response to Electroconvulsive Therapy via Machine Learning on Structural Magnetic Resonance Imaging Data. *JAMA Psychiatry*. 2016; 73: 557–564. <https://doi.org/10.1001/jamapsychiatry.2016.0316>.
- [13] Liu XP, Jin X, Seyed Ahmadian S, Yang X, Tian SF, Cai YX, *et al.* Clinical significance and molecular annotation of cellular morphometric subtypes in lower-grade gliomas discovered by machine learning. *Neuro-oncology*. 2023; 25: 68–81. <https://doi.org/10.1093/neuonc/noac154>.
- [14] Xi C, Lu B, Guo X, Qin Z, Yan C, Hu S. Characteristics of brain network connectome and connectome-based efficacy predictive model in bipolar depression. *Molecular Psychiatry*. 2025; 30: 5150–5160. <https://doi.org/10.1038/s41380-025-03099-6>.

- [15] Gao Y, Guo X, Zhong Y, Liu X, Tian S, Deng J, *et al*. Decreased dorsal attention network homogeneity as a potential neuroimaging biomarker for major depressive disorder. *Journal of Affective Disorders*. 2023; 332: 136–142. <https://doi.org/10.1016/j.jad.2023.03.080>.
- [16] Brinkman L, Todorov A, Dotsch R. Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*. 2017; 28: 333–361. <https://doi.org/10.1080/10463283.2017.1381469>.
- [17] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology*. 1999; 56: 303–308. <https://doi.org/10.1001/archneur.56.3.303>.
- [18] Irish M, Piguet O, Hodges JR. Self-projection and the default network in frontotemporal dementia. *Nature Reviews. Neurology*. 2012; 8: 152–161. <https://doi.org/10.1038/nrneuro.2012.11>.
- [19] Oppedal K, Ferreira D, Cavallin L, Lemstra AW, Ten Kate M, Padovani A, *et al*. A signature pattern of cortical atrophy in dementia with Lewy bodies: A study on 333 patients from the European DLB consortium. *Alzheimer's & Dementia: the Journal of the Alzheimer's Association*. 2019; 15: 400–409. <https://doi.org/10.1016/j.jalz.2018.09.011>.
- [20] Hainmueller T, Bartos M. Dentate gyrus circuits for encoding, retrieval and discrimination of episodic memories. *Nature Reviews. Neuroscience*. 2020; 21: 153–168. <https://doi.org/10.1038/s41583-019-0260-z>.
- [21] Kitamura T, Ogawa SK, Roy DS, Okuyama T, Morrissey MD, Smith LM, *et al*. Engrams and circuits crucial for systems consolidation of a memory. *Science (New York, N.Y.)*. 2017; 356: 73–78. <https://doi.org/10.1126/science.aam6808>.
- [22] Connor CE, Knierim JJ. Integration of objects and space in perception and memory. *Nature Neuroscience*. 2017; 20: 1493–1503. <https://doi.org/10.1038/nn.4657>.
- [23] Salvador R, Suckling J, Coleman MR, Pickard JD, Menon D, Bullmore E. Neurophysiological architecture of functional magnetic resonance images of human brain. *Cerebral Cortex (New York, N.Y.: 1991)*. 2005; 15: 1332–1342. <https://doi.org/10.1093/cercor/bhi016>.
- [24] Sporns O, Tononi G, Kötter R. The human connectome: A structural description of the human brain. *PLoS Computational Biology*. 2005; 1: e42. <https://doi.org/10.1371/journal.pcbi.0010042>.
- [25] Li X, Steffens DC, Potter GG, Guo H, Song S, Wang L. Decreased between-hemisphere connectivity strength and network efficiency in geriatric depression. *Human Brain Mapping*. 2017; 38: 53–67. <https://doi.org/10.1002/hbm.23343>.