Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Task-evoked multimodal wearable sensing for community-based cognitive impairment assessment

Chengyu Liu

chengyu@seu.edu.cn

State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University https://orcid.org/0000-0003-1965-3020

Minghui Zhao

State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University https://orcid.org/0009-0002-4585-7107

Lulu Zhao

State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University

Haiyan Yin

School of Nursing, Nanjing University of Chinese Medicine

Yulei Song

School of Nursing, Nanjing University of Chinese Medicine

Xinru Qi

School of Nursing, Nanjing University of Chinese Medicine

Wentao Xiang

Engineering Research Center of Intelligent Theranostics Technology and Instruments, Ministry of Education, School of Biomedical Engineering and Informatics, Nanjing Medical University

Chenxi Yang

State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University

Yi Zhu

Department of Rehabilitation Medicine and Department of Radiology, The First Affiliated Hospital of Nanjing Medical University

Yu Liu

University of Oxford https://orcid.org/0000-0002-2399-2829

Hongxiang Gao

State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University

Jianpeng Gu
School of Nursing, Nanjing University of Chinese Medicine
Yifan Fei
State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University
Jiexin Zhang
State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University
Lukai Pang
State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University
Hanbo Li
Department of Rehabilitation Medicine and Department of Radiology, The First Affiliated Hospital of Nanjing Medical University
Shui Tian
Department of Rehabilitation Medicine and Department of Radiology, The First Affiliated Hospital of Nanjing Medical University
Yantao Xing
State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University
Heyang Xu
State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University
Jianqing Li
State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University
Yamei Bai
School of Nursing, Nanjing University of Chinese Medicine
Tingting Zhu
University of Oxford https://orcid.org/0000-0002-1552-5630

Article

Keywords: Cognitive impairment assessment, Multimodal sensing system, Physiological and behavioral signals, Multidimensional cognitive functions, Cognitive tasks

Posted Date: August 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8946251/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Task-evoked multimodal wearable sensing for community-based cognitive impairment assessment

Minghui Zhao[1†], Lulu Zhao[1†], Haiyan Yin[2†], Yulei Song[2†], Xinru Qi[2], Wentao Xiang[3], Chenxi Yang[1], Yi Zhu[4], Yu Liu[5], Hongxiang Gao[1], Jianpeng Gu[2], Yifan Fei[1], Jiexin Zhang[1], Lukai Pang[1], Hanbo Li[4], Shui Tian[4], Yantao Xing[1], Heyang Xu[1], Jianqing Li[1*], Yamei Bai[2*], Tingting Zhu[5*], Chengyu Liu[1*]

[1*]State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University, Nanjing, 210096, China.
[2]School of Nursing, Nanjing University of Chinese Medicine, Nanjing, 210023, China.
[3]Engineering Research Center of Intelligent Theranostics Technology and Instruments, Ministry of Education, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, 211166, China.
[4]Department of Rehabilitation Medicine and Department of Radiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, 210029, China.
[5]Department of Engineering Science, University of Oxford, Oxford, OX3 7DQ, UK.

*Corresponding author(s). E-mail(s): ljq@seu.edu.cn; czbym@njucm.edu.cn; tingting.zhu@eng.ox.ac.uk; chengyu@seu.edu.cn;
Contributing authors: zmh56@seu.edu.cn; zhaolulu@seu.edu.cn; 370551@njucm.edu.cn; songyulei@njucm.edu.cn; 20221081@njucm.edu.cn; xiangbmu@njmu.edu.cn; chenxiyang@seu.edu.cn; ZhuYi1981@njmu.edu.cn; yu.liu@eng.ox.ac.uk; hongxiang_seu@seu.edu.cn; 20230980@njucm.edu.cn; 220243742@seu.edu.cn; 220243743@seu.edu.cn; 230239011@seu.edu.cn; lihanbo@stu.njmu.edu.cn; shuitian1590@njmu.edu.cn; xingyt@seu.edu.cn; xuhy@seu.edu.cn;
[†]These authors contributed equally to this work.

Abstract

With global population aging, cognitive impairment has become a major public health challenge. Current assessment approaches provide important information on cognitive function and disease-related pathology, but offer limited characterization of task-evoked physiological and behavioral responses associated with cognitive impairment. Here, we investigated task-evoked physiological and behavioral responses using a wearable sensing system that synchronously records electroencephalography, electrocardiography, speech, and video signals during structured cognitive tasks. A total of 1,385 community-dwelling adults aged 60 years or older were recruited from five administrative districts in Nanjing, China, and divided into an internal development cohort (n=831) and an external validation cohort (n=554). The system distinguished normal cognition from cognitive impairment, with areas under the receiver operating characteristic curve of 0.93 and 0.90 in the internal and external cohorts, respectively. These findings demonstrate the utility of wearable multimodal signals for accurate assessment of cognitive impairment. Importantly, different cognitive tasks elicit distinct physiological-behavioral features that contribute to cognitive impairment assessment and reflect distinct, complementary cognitive functions. These findings suggest that task-evoked wearable multimodal signals can provide process-level functional phenotyping information for cognitive impairment assessment, complementing conventional cognitive assessment.

Keywords: Cognitive impairment assessment; Multimodal sensing system; Physiological and behavioral signals; Multidimensional cognitive functions; Cognitive tasks

1 Introduction

As the global population ages rapidly, cognitive impairment has become a major public health challenge [1]. Cognitive impairment, ranging from mild cognitive impairment to dementia, is often characterized by multidomain functional abnormalities and neuropsychiatric symptoms [2, 3]. Currently, more than 57 million people worldwide are living with dementia, with over 60% residing in low- and middle-income countries, and nearly 10 million new cases diagnosed each year [1]. By 2050, the global number of dementia cases is projected to reach 152 million [4], highlighting the growing burden of cognitive impairment on individuals, families, and healthcare systems worldwide [5].

In the mild cognitive impairment stage, clinical manifestations are often subtle and easily attributed to normal aging, making timely identification challenging in community and routine clinical settings [6]. This highlights the need for multidimensional functional phenotyping beyond single categorical judgments [5]. Cognitive impairment may involve heterogeneous abnormalities across memory, language, attention, visuospatial/executive ability, and orientation [7, 8], with manifestations influenced by mental health, aging-related physiological changes, education, and daily functioning [9, 10]. Such characterization can support clinical interpretation and intervention planning [11]. However, how cognitive impairment is reflected in physiological and behavioral responses during cognitive task performance remains insufficiently characterized.

Existing assessment approaches provide important information on cognitive function and disease-related pathology [12, 13]. Neuropsychological testing remains the standard for evaluating cognitive functions [14, 15], but scale scores provide limited characterization of the physiological and behavioral processes underlying task performance. Neuroimaging and cerebrospinal fluid biomarkers characterize brain structure, function, and molecular pathology [16, 17]. Blood-based biomarkers are advancing for Alzheimer's disease-related amyloid and tau pathology [18, 19], while genetic testing may inform disease susceptibility in specific contexts [20]. Objective functional phenotyping based on task-evoked physiological and behavioral responses may therefore provide additional information for cognitive impairment assessment.

Multimodal sensing [21, 22] provides a non-invasive framework for characterizing task-evoked physiological and behavioral responses by synchronously capturing signals such as electroencephalography (EEG), electrocardiography (ECG), speech, and video [23, 24]. These modalities can reflect different aspects of cognitive performance. EEG-based features provide information on task-related neural activity and spectral changes [25, 26]; ECG-based features, particularly heart rate variability, may reflect autonomic regulation during cognitive load [27, 28]; speech-based acoustic and semantic features can characterize language production and cognitive organization [29, 30]; and video-based behavioral features may capture facial expression, attention, and task-related observable responses [31, 32]. Integrating these modalities may therefore provide a more comprehensive description of how neural activity, autonomic regulation, language production, and observable behavior jointly vary with cognitive function and cognitive impairment status.

Existing digital biomarker studies have examined cognitive impairment using different sensing modalities, including actigraphy [33], speech [34, 35], GPS-based mobility or driving behavior [36], wearable EEG [25], and ECG-derived heart rate variability [37]. Multimodal passive sensing systems have further integrated ambient and wearable sensors with federated learning to monitor digital biomarkers in naturalistic living environments [38]. These studies indicate that physiological and behavioral signals can capture cognitive-related information. However, prior work has mainly focused on isolated modalities, passive monitoring, or naturalistic settings, with limited characterization of coordinated multimodal responses during structured cognitive task execution. The present study therefore investigates task-evoked physiological and behavioral responses and evaluates their associations with cognitive function and cognitive impairment status.

To address this gap, we developed a wearable multimodal sensing system to investigate task-evoked physiological and behavioral responses during structured tasks targeting seven core cognitive functions [39]. Using a prospective clinical cohort, we examined whether these multimodal responses are associated with cognitive function and cognitive impairment status. We further evaluated the contribution of different tasks and modalities, and mapped multimodal signal patterns to specific cognitive functions. This framework provides objective, process-level functional phenotyping information that

complements neuropsychological testing and pathological biomarkers by characterizing how cognitive impairment is reflected in task-evoked physiological and behavioral responses.

2 Results

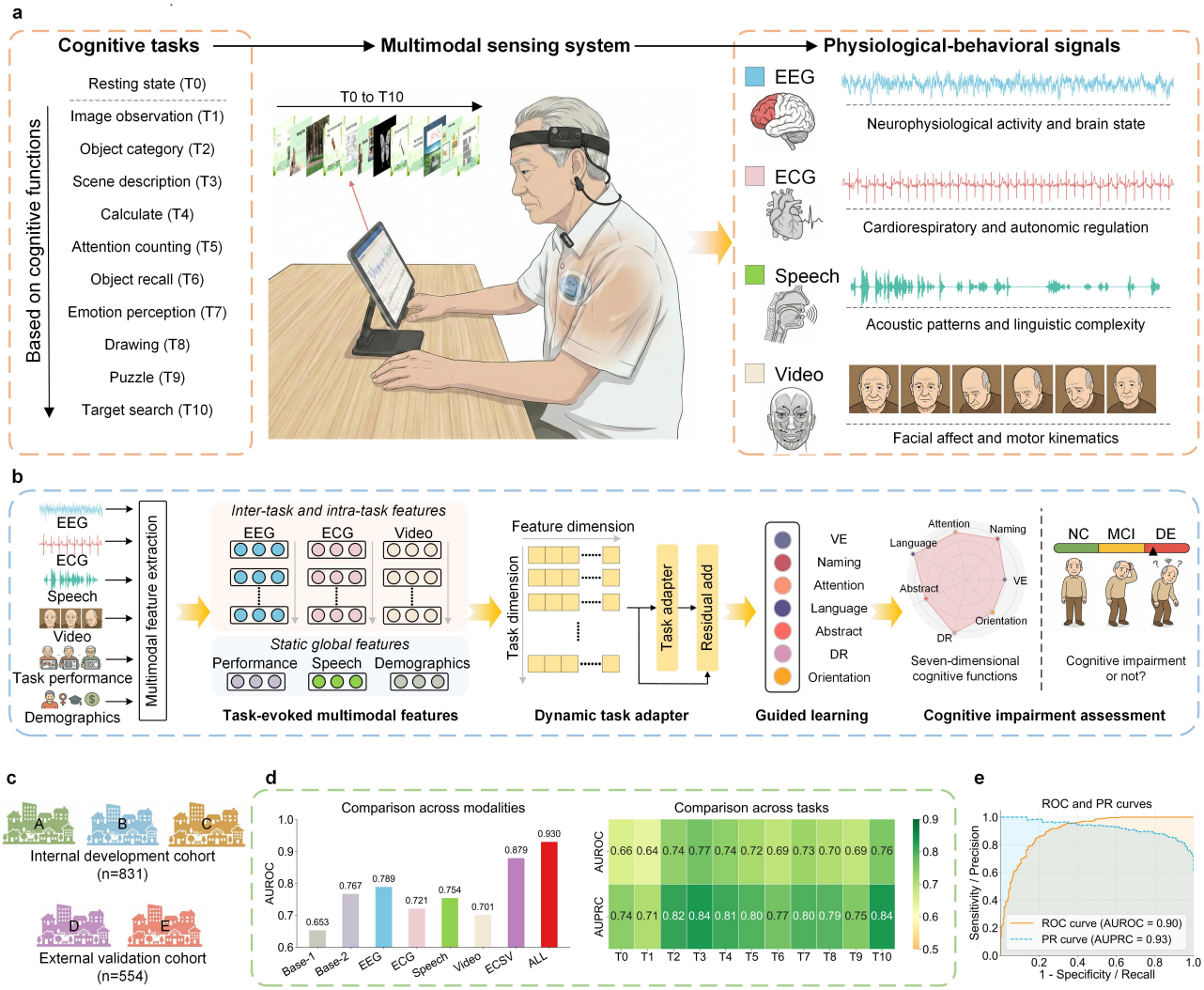


Fig. 1 Multimodal sensing system based on wearable devices for cognitive impairment assessment. a, The multimodal system integrates a prefrontal EEG headband, wearable ECG patch, microphone, and camera to collect synchronized physiological and behavioral signals. The experimental protocol consists of a resting-state period (T0) followed by a series of active cognitive tasks (T1–T10). These tasks are designed to evaluate specific cognitive functions, including visuospatial/executive (VE), naming, attention, language, abstraction, delayed recall (DR), and orientation. **b**, The proposed model employs a dynamic task adapter with learning guided by cognitive functions for cognitive impairment assessment. It integrates task-evoked physiological and behavioral signals (EEG, ECG, speech, and video) with task performance metrics (scores and completion time) and demographic information. **c**, Construction of internal development and external validation cohorts via multi-district data collection. **d**, Comparative assessment of cognitive impairment in the internal development cohort. Base-1, Base-2, ECSV, and ALL denote different feature sets: Base-1 represents baseline results from demographic information; Base-2 represents baseline results from task performance; ECSV integrates EEG, ECG, speech, and video features; ALL combines all modalities. **e**, Generalization capability in the external validation cohort illustrated by Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves.

We developed a wearable sensing system that synchronously captures multimodal signals during structured cognitive tasks (T0–T10) in Fig. 1a. The system records EEG, ECG, speech, and video signals to characterize physiological and behavioral responses evoked by cognitive tasks (Supplementary Fig. 1). We then summarized these signals into task-related features and used a dynamic task adapter model to estimate seven cognitive dimensions, supporting an interpretable assessment of cognitive impairment (Fig. 1b and Supplementary Fig. 2). We validated this system in a prospective clinical cohort of 1,385 community-dwelling older adults in Nanjing, which was stratified into an internal development cohort recruited from three districts (Qixia, Yuhuatai, and Pukou) and a geographically independent external

validation cohort from distinct districts (Gulou and Jiangning) in Fig. 1c. This cohort design allowed us to examine task-related signal patterns in the development cohort and to test the generalizability of the assessment system in an independent external cohort (Fig. 1d,e and Supplementary Fig. 6).

2.1 Cohort description and statistical summary

Table 1 Demographic and clinical characteristics of the study participants

	Total (n=1,385)	NC (n=540)	Cognitive Impairment (n=845)	
			MCI (n=777)	DE (n=68)
Internal development cohort				
Participants, n	831	324	464	43
Sex (female, male, %)	(68.7, 31.3)	(71.6, 28.4)	(69.2, 30.8)	(41.9, 58.1)
Age	71.53±7.51	70.11±7.22	71.50±6.75	82.53±8.44
Education	8.12±3.67	8.63±3.82	7.61±3.42	9.90±4.14
MoCA	21.00±4.88	24.97±2.54	18.98±3.69	12.71±6.05
VE	2.35±1.32	3.15±1.23	1.93±1.06	0.88±1.04
Naming	2.50±0.78	2.84±0.43	2.33±0.82	1.90±1.27
Attention	5.01±1.21	5.65±0.64	4.73±1.20	3.19±1.69
Language	2.10±0.86	2.56±0.65	1.84±0.83	1.43±0.94
Abstract	1.01±0.72	1.35±0.64	0.79±0.69	0.74±0.73
DR	2.42±1.59	3.57±1.11	1.76±1.39	0.83±1.19
Orientation	5.57±0.99	5.85±0.43	5.60±0.81	3.19±2.07
MMSE	25.28±3.83	27.63±1.90	24.29±3.34	17.91±5.88
SCD-Q9	4.42±2.11	4.18±1.98	4.45±2.15	5.95±2.09
HAMA	4.22±3.20	3.29±3.04	4.76±3.18	5.36±2.99
GDS	0.90±1.47	0.69±1.28	1.05±1.54	0.98±1.88
HIS	2.22±1.13	2.17±1.02	2.30±1.18	1.86±1.35
ADL	21.71±6.70	20.27±0.79	20.48±1.06	46.33±15.35
FRAIL	0.42±0.76	0.28±0.57	0.35±0.67	2.14±0.89
External validation cohort				
Participants, n	554	216	313	25
Sex (female, male, %)	(65.9, 34.1)	(63.9, 36.1)	(68.1, 31.9)	(56.0, 44.0)
Age	71.43±7.40	70.05±6.59	71.52±7.24	82.12±7.28
Education	7.83±3.65	8.38±3.70	7.46±3.42	7.60±5.19
MoCA	20.84±4.90	24.84±2.47	18.73±3.90	12.56±5.29
VE	2.31±1.31	2.96±1.22	1.97±1.16	0.88±1.05
Naming	2.54±0.72	2.82±0.42	2.39±0.78	1.84±1.03
Attention	4.95±1.23	5.59±0.65	4.62±1.30	3.60±1.41
Language	2.02±0.90	2.56±0.63	1.71±0.86	1.16±0.85
Abstract	0.99±0.74	1.36±0.63	0.79±0.70	0.36±0.64
DR	2.41±1.61	3.61±1.16	1.71±1.36	0.72±1.24
Orientation	5.61±0.94	5.93±0.30	5.55±0.88	3.68±2.21
MMSE	25.20±3.83	27.53±1.95	24.11±3.54	18.68±5.63
SCD-Q9	4.52±1.99	4.16±1.78	4.73±2.06	5.08±2.47
HAMA	4.33±3.46	2.87±2.66	5.13±3.48	6.92±4.57
GDS	0.94±1.63	0.55±1.08	1.14±1.78	1.76±2.68
HIS	2.19±1.08	1.97±0.95	2.33±1.12	2.20±1.26
ADL	21.35±4.89	20.31±0.90	20.56±1.28	40.24±11.56
FRAIL	0.39±0.78	0.19±0.52	0.36±0.69	2.40±0.96

NC, normal cognition; MCI, mild cognitive impairment; DE, dementia; MoCA, Montreal Cognitive Assessment; MMSE, Mini-Mental State Examination; VE, visuospatial/executive; DR, delayed recall; SCD-Q9, Subjective Cognitive Decline Questionnaire-9; HAMA, Hamilton Anxiety Rating Scale; GDS, Geriatric Depression Scale; HIS, Hachinski Ischemic Score; ADL, Activities of Daily Living Scale; FRAIL, Frailty Assessment Scale. Education reflects years of formal schooling completed. Values are mean ± s.d. unless otherwise specified.

A total of 1,608 community-dwelling older adults aged 60 years and above were recruited for the study. After applying predefined inclusion and exclusion criteria, 1,385 participants were included in

the final analysis. The cohort included 540 individuals with normal cognition (NC), 777 with mild cognitive impairment (MCI), and 68 with dementia (DE). Clinical cognitive status was adjudicated by neurologists based on comprehensive standard assessment criteria integrating cognitive and functional assessments [5], with the detailed grouping procedure provided in Supplementary Fig. 28. Based on geographic location and time of collection, the cohort was divided into two parts: an internal development cohort (n = 831) and an independent external validation cohort (n = 554). The detailed participant inclusion flowchart is provided in Supplementary Fig. 6 and 27.

Internal development cohort: Demographic characteristics and the distributions of cognitive, emotional, and functional assessments are presented in Table 1. With increasing cognitive impairment, participants were progressively older (NC: 70.11 ± 7.22 years; MCI: 71.50 ± 6.75 years; DE: 82.53 ± 8.44 years). Years of education differed across groups (NC: 8.63 ± 3.82 years; MCI: 7.61 ± 3.42 years; DE: 9.90 ± 4.14 years), reflecting heterogeneity that may influence cognitive trajectories.

Both the Montreal Cognitive Assessment (MoCA) [39] and Mini-Mental State Examination (MMSE) [40] scores differed across the three groups, with stepwise declines from NC to MCI and DE. The average MoCA score in the NC group was 24.97 ± 2.54, which was higher than in the MCI (18.98 ± 3.69) and DE groups (12.71 ± 6.05). Similarly, the MMSE scores were 27.63 ± 1.90 for NC, 24.29 ± 3.34 for MCI, and 17.91 ± 5.88 for DE. Scores across all seven cognitive functions, including visuospatial/executive, naming, attention, language, abstraction, delayed recall, and orientation, declined progressively, with the steepest decrease in the DE group. For subjective cognitive decline, Subjective Cognitive Decline Questionnaire-9 (SCD-Q9) scores were higher in the DE group (5.95 ± 2.09) than in the MCI (4.45 ± 2.15) and NC groups (4.18 ± 1.98), indicating more frequent self- or informant-reported cognitive concerns.

Cognitive, emotional, and functional measures jointly reflect the multidimensional aspects of mental health in aging populations. Regarding emotional measures [5], the MCI and DE groups showed higher scores on both the Hamilton Anxiety Rating Scale (HAMA) and Geriatric Depression Scale (GDS) compared to the NC group, suggesting that anxiety and depressive symptoms are more common among individuals with cognitive impairment. Scores on the Activities of Daily Living (ADL) scale, where higher scores indicate greater impairment, were also significantly elevated in the DE group, reflecting reduced daily functioning. Frailty Assessment Scale (FRAIL) scores similarly showed a progressive increase with worsening cognitive status.

External validation cohort: This cohort included 554 participants, comprising 216 NC, 313 MCI, and 25 DE individuals, as shown in Table 1. Compared with the development cohort, there were slight demographic differences, but overall age and education distributions were broadly similar. Cognitive assessments, including MoCA and MMSE, as well as domain-specific scores, showed stepwise declines from NC to MCI and DE, consistent with the development cohort. These consistent trends across two independent cohorts, despite minor demographic variations, provided a stable basis for evaluating the proposed assessment approach (detailed cohort distributions are provided in Supplementary Figs. 7–12).

2.2 Task-evoked multimodal features characterize cognitive impairment

We performed unified preprocessing and feature extraction on synchronously acquired EEG, ECG, speech, and video signals. A set of physiological and behavioral features were extracted for subsequent modeling and interpretability analysis, with detailed feature definitions provided in Supplementary Tables 14–18. To systematically characterize the relationship between these multimodal signatures and cognitive impairment, we analyzed the correlation patterns between task-evoked features and cognitive functions across different tasks (Fig. 2).

Intra-task analysis highlights distinct, task-specific patterns in multimodal signal contributions (Fig. 2a). The detailed associations between specific features and cognitive impairment are presented in Supplementary Figs. 3, 4. EEG and speech modalities exhibit the most consistent and dense connections with high-order cognitive functions, particularly executive function, attention, and language (Supplementary Fig. 21). Specifically, nonlinear complexity and information entropy derived from EEG and ECG were associated with visuospatial/executive functions and attention, indicating that task-evoked physiological responses may reflect altered neural and autonomic regulation under cognitive load. Speech-derived prosodic and semantic features were related to language and memory domains, while video-based facial kinematics, facial geometry, and eye movement patterns provided complementary behavioral information related to emotion regulation and abstraction.

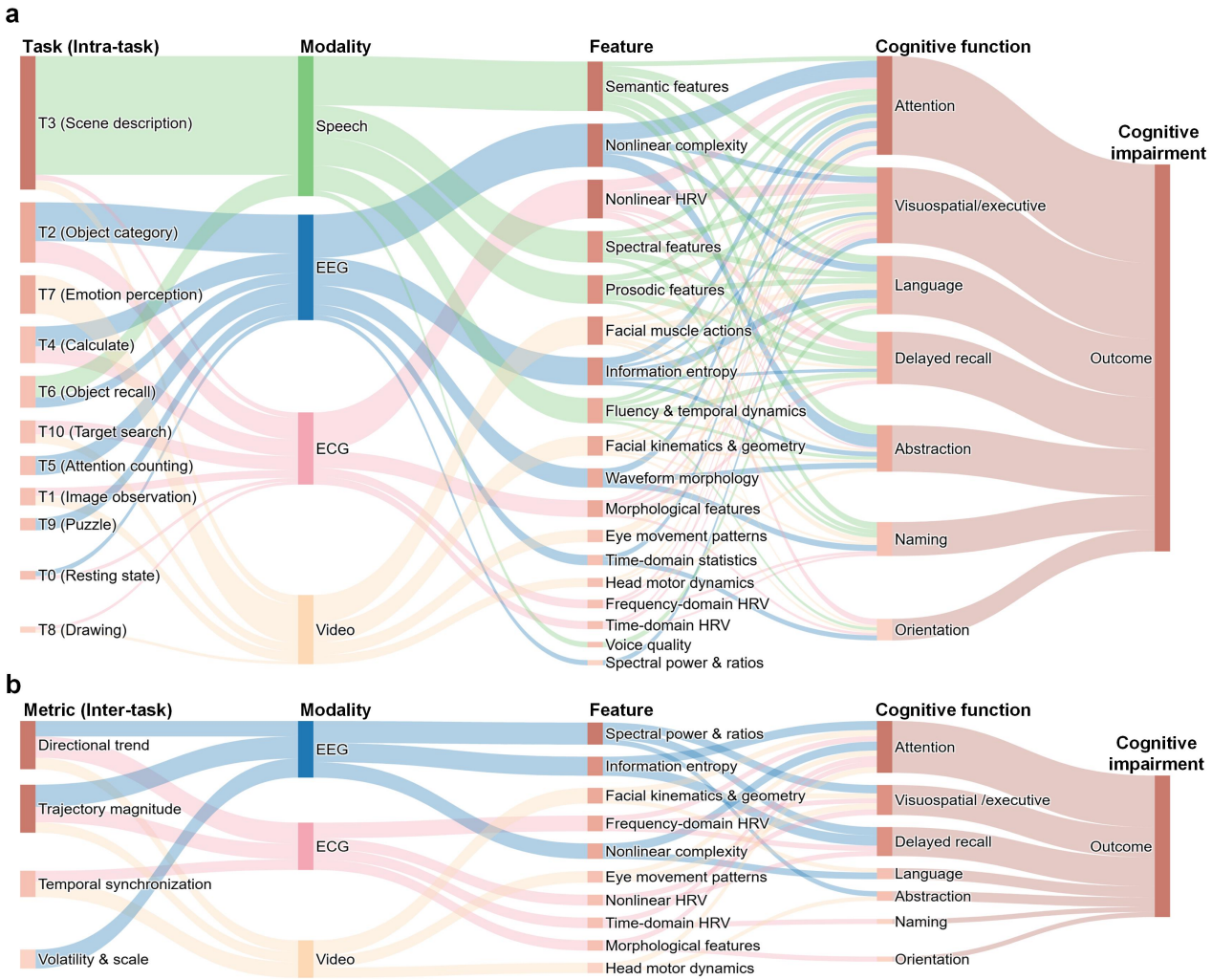


Fig. 2 Mapping task-evoked multimodal features to cognitive functions and impairment status. a, Intra-task multimodal feature associations. Alluvial diagrams map the flow from cognitive tasks (left) through sensing modalities (speech, EEG, ECG, video) and extracted feature sets to seven cognitive functions (derived from MoCA). This panel highlights specific task-modality combinations that explain variations in specific cognitive functions. **b**, Inter-task dynamic feature associations. This panel illustrates how metrics capturing feature evolution over time, specifically trajectory, directional trend, and volatility throughout the assessment, correlate with cognitive performance. In both panels, flow width is proportional to the aggregated absolute Spearman correlation strength (FDR-corrected) between source and target nodes; thicker flows indicate stronger correlations between specific feature sets and cognitive functions.

Beyond intra-task features, we analyzed inter-task feature evolution across the assessment timeline to capture subject adaptability. Fig. 2b visualizes the impact of these dynamic metrics. For detailed statistical metrics of the trajectories across tasks, please refer to Supplementary Table 18. We observed that the trajectory magnitude and directional trends of EEG spectral power and ECG heart rate variability (HRV) are closely correlated with sustained attention and delayed recall. This suggests that both the magnitude and direction of physiological changes during testing may provide complementary information beyond absolute feature values (Supplementary Fig. 22). Additionally, the integrated evolution of facial kinematics, eye movement patterns, and head motor dynamics correlates with executive function and attention, suggesting that observable motor instability serves as a robust behavioral marker for cognitive impairment.

2.3 Multimodal fusion improves assessment performance and robustness

In the internal development cohort, we evaluated whether multimodal task-evoked physiological and behavioral signals could distinguish individuals with cognitive impairment (MCI/DE) from cognitively normal (NC) individuals. As shown in Table 2, we compared demographic information, task performance, individual sensor modalities, and their multimodal combinations. The combination of

Table 2 Comparison of assessment performance across modalities and cognitive tasks

	AUROC (95% CI)	AUPRC (95% CI)	Accuracy (95% CI)
Modality[a]			
Demographic	0.653 (0.616–0.691)	0.750 (0.709–0.790)	0.584 (0.549–0.617)
Task performance	0.767 (0.735–0.799)	0.843 (0.810–0.871)	0.696 (0.663–0.727)
EEG(E)	0.789 (0.758–0.820)	0.854 (0.821–0.882)	0.730 (0.699–0.760)
ECG(C)	0.721 (0.686–0.757)	0.787 (0.746–0.826)	0.688 (0.656–0.720)
Speech(S)	0.754 (0.721–0.786)	0.831 (0.797–0.863)	0.665 (0.632–0.698)
Video(V)	0.701 (0.665–0.736)	0.798 (0.762–0.834)	0.605 (0.571–0.639)
ECSV	0.879 (0.854–0.903)	0.903 (0.875–0.930)	0.819 (0.792–0.845)
Demographic+Performance+ECSV	0.930 (0.913–0.947)	0.956 (0.943–0.968)	0.858 (0.832–0.881)
Task[b]			
T0 (Resting state)	0.659 (0.621–0.696)	0.744 (0.699–0.786)	0.614 (0.580–0.647)
T1 (Image observation)	0.639 (0.600–0.677)	0.708 (0.664–0.755)	0.616 (0.582–0.649)
T2 (Object category)	0.740 (0.706–0.773)	0.820 (0.786–0.853)	0.684 (0.651–0.715)
T3 (Scene description)	0.767 (0.735–0.800)	0.838 (0.803–0.868)	0.708 (0.675–0.738)
T4 (Calculate)	0.741 (0.707–0.774)	0.811 (0.773–0.846)	0.668 (0.635–0.700)
T5 (Attention counting)	0.717 (0.682–0.752)	0.796 (0.757–0.833)	0.685 (0.652–0.716)
T6 (Object recall)	0.692 (0.656–0.729)	0.768 (0.727–0.807)	0.676 (0.643–0.708)
T7 (Emotion perception)	0.727 (0.692–0.761)	0.798 (0.759–0.836)	0.662 (0.629–0.694)
T8 (Drawing)	0.704 (0.668–0.739)	0.793 (0.758–0.828)	0.663 (0.630–0.695)
T9 (Puzzle)	0.689 (0.653–0.726)	0.753 (0.709–0.795)	0.658 (0.625–0.690)
T10 (Target search)	0.755 (0.723–0.788)	0.836 (0.804–0.865)	0.696 (0.663–0.727)
Neuropsychological test[c]			
MMSE	0.836 (0.809–0.862)	0.876 (0.851–0.899)	0.777 (0.747–0.805)
MoCA	0.922 (0.904–0.938)	0.942 (0.927–0.954)	0.884 (0.861–0.905)

All performance metrics were obtained via stratified 5-fold cross-validation on the internal development dataset and are reported as the micro-average across all folds [16]. All metrics are reported as decimals (0-1) with 95% confidence intervals (CI) [12]. The 95% CI of area under the receiver operating characteristic curve (AUROC) was computed using the non-parametric method of DeLong [41]. The 95% CI of area under the precision–recall curve (AUPRC) was computed using 2,000 bootstrap resampling. The 95% CI of accuracy was obtained using the Clopper-Pearson method. Detailed cross-modal and task-specific comparison results are provided in Supplementary Tables 2 and 3.

[a]Task performance refers to participants' scores and time during cognitive tasks; ECSV represents integrated features from four modalities (EEG, ECG, speech, and video).

[b]T0 represents the resting-state acquisition. T1–T10 correspond to ten specific tasks designed based on multidimensional cognitive functions.

[c]The performance metrics of MMSE and MoCA are likely overestimated as they informed the clinical assessment; these results serve as reference benchmarks.

demographic information, task performance, and multimodal physiological-behavioral signals achieved the best overall performance, with an AUROC of 0.930 (95% CI: 0.913-0.947). These results indicate that sensor-derived features provide additional information beyond demographic and task-performance measures alone (see Supplementary Table 2 and Fig. 13 for details).

Among the individual modalities, EEG showed the reasonable discrimination ability (AUROC = 0.789, 95% CI: 0.758-0.820), suggesting that task-evoked neural activity was closely related to cognitive status. Speech features also showed informative performance (AUROC = 0.754, 95% CI: 0.721-0.786), indicating that language production and acoustic characteristics may reflect cognitive impairment. ECG and video features showed moderate but complementary discrimination ability. Importantly, the fused ECSV model performed better than any single modality, supporting the value of integrating neural, autonomic, speech, and facial behavioral responses during structured cognitive tasks.

Further analysis of the resting state and ten task conditions (T0–T10) revealed task-specific differences in physiological and behavioral signal responses, as shown in Table 2 (see Supplementary Table 3 and Fig. 14 for details). Tasks T3 (scene description) and T10 (target search) achieved relatively higher AUROC of 0.767 (95% CI: 0.735–0.800) and 0.755 (95% CI: 0.723–0.788), respectively. Notably, T3, T10, T4 (calculation) and T2 (object categorization) also yielded higher AUPRC, indicating their particular sensitivity in detecting cognitive impairment.

Inter-task feature interactions proved critical for model performance. While static single-task features showed effective discriminative power, modeling temporal dynamics across task transitions enhanced informativeness (Supplementary Fig. 5a-c). This sequential approach effectively captured physiological and behavioral variations, ranging from EEG fluctuations and autonomic regulation in ECG to facial action shifts across emotional valences (Supplementary Fig. 22). Extensive comparative analyses confirmed that our model outperforms established benchmarks (Supplementary Figs. 15–18 and Supplementary Tables 4–7).

2.4 Robust generalization to external cohorts under readjusted prevalence

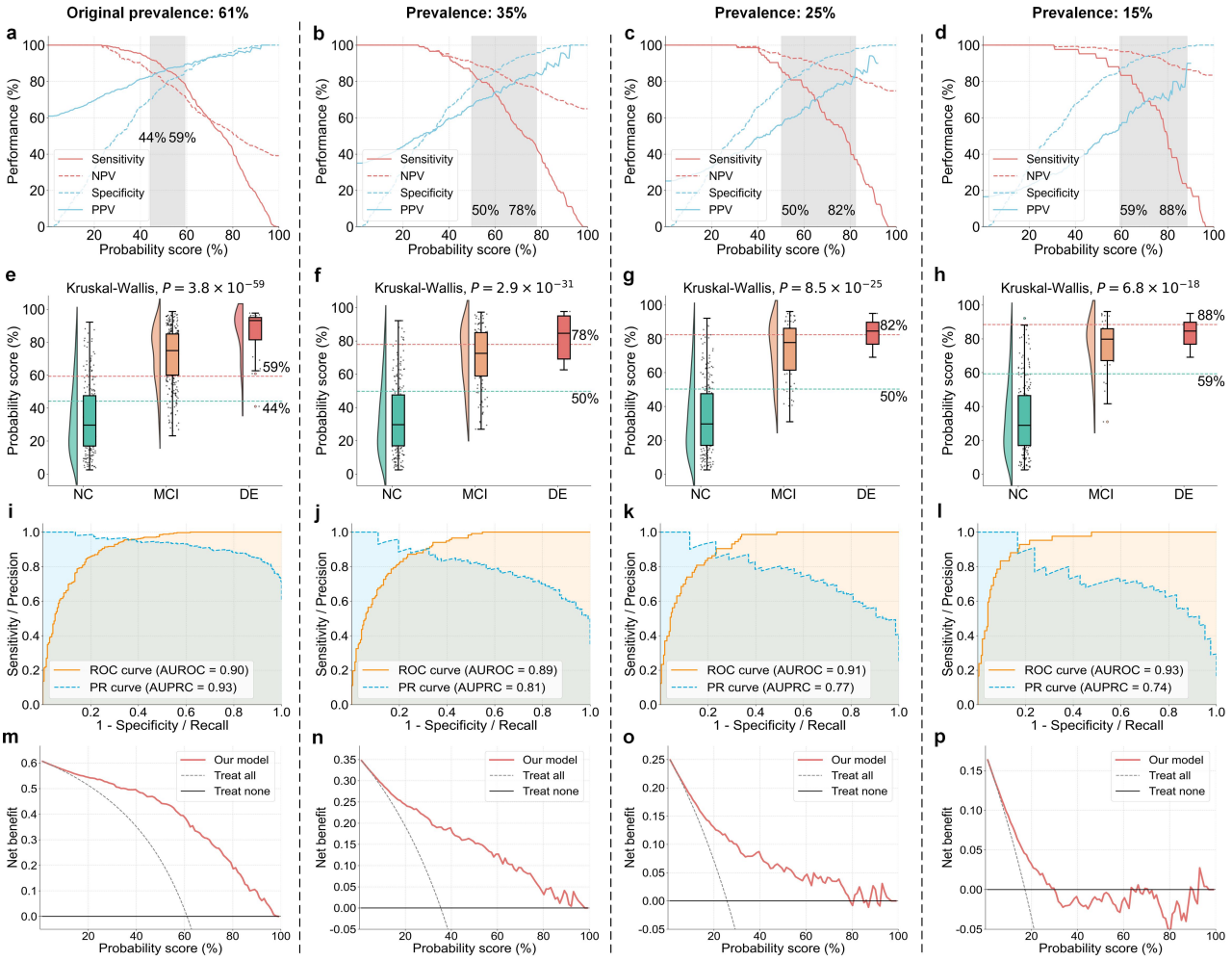


Fig. 3 Comparison of model performance under readjusted cognitive impairment prevalence. This figure demonstrates the robustness of the two-cutoff probability stratification strategy across three cognitive states under four prevalence settings: original distribution (**a, e, i, m**), 35% (**b, f, j, n**), 25% (**c, g, k, o**), and 15% (**d, h, l, p**). **a–d**, Sensitivity (solid red), specificity (solid blue), PPV (dashed blue), and NPV (dashed red) as functions of probability score across four prevalence settings. The shaded gray area marks the intermediate range (Performance $\geq$ 85%), bounded by low-probability and high-probability thresholds. **e–h**, Probability score distributions across cognitive groups using violin and box plots, with dashed lines indicating classification cutoffs. Significant group differences are confirmed by Kruskal–Wallis tests [42] (all $P < 0.001$). **i–l**, Receiver operating characteristic (ROC) and precision–recall (PR) curves across prevalence settings, illustrating consistent discriminative ability of the model. **m–p**, Decision Curve Analysis (DCA) across continuous threshold probabilities, showing the net clinical benefit of the model relative to the "treat-all" and "treat-none" reference strategies under each prevalence setting.

To further assess model generalizability, we performed external validation on an independently collected cohort across four prevalence settings [5]: the original cohort ratio and three readjusted scenarios with cognitively impaired proportions set to 35%, 25%, and 15%. A two-cutoff stratification strategy [12] was applied to divide model output probabilities into low-, intermediate-, and high-probability zones. Fig. 3a-d shows sensitivity, specificity, PPV, and NPV across thresholds, with the shaded intermediate-probability range bounded by a lower cutoff (sensitivity and NPV $\geq$ 85%) and an upper

cutoff (specificity and PPV $\geq 85\%$). Under the original prevalence, the lower cutoff supported rule-out and the upper cutoff supported identification of high-probability cases; under readjusted prevalence settings, the cutoff boundaries shifted accordingly, indicating that this strategy should be interpreted in a prevalence-aware manner. Fig. 3e-h shows the distribution of model-derived probabilities across NC, MCI, and DE participants. NC participants were mainly located in the low-probability range, DE participants were concentrated in the high-probability range, and MCI participants showed an intermediate distribution.

Fig. 3i-l presents the ROC and precision-recall curves, showing that the model preserved discrimination across different prevalence settings. Finally, Fig. 3m-p presents the Decision Curve Analysis (DCA). The model showed positive net benefit under the original 61%, 35%, and 25% prevalence settings. Although the range of positive net benefit became narrower under the 15% low-prevalence setting, the model retained clinical benefit at lower threshold probabilities, supporting its potential role in front-line community screening where sensitive identification of possible cognitive impairment cases is prioritized. This limitation at higher threshold probabilities is partly attributable to the mathematically amplified sensitivity of DCA to false positives when positive cases are extremely sparse ($n = 42$).

2.5 Subgroup analyses and robustness of model discrimination

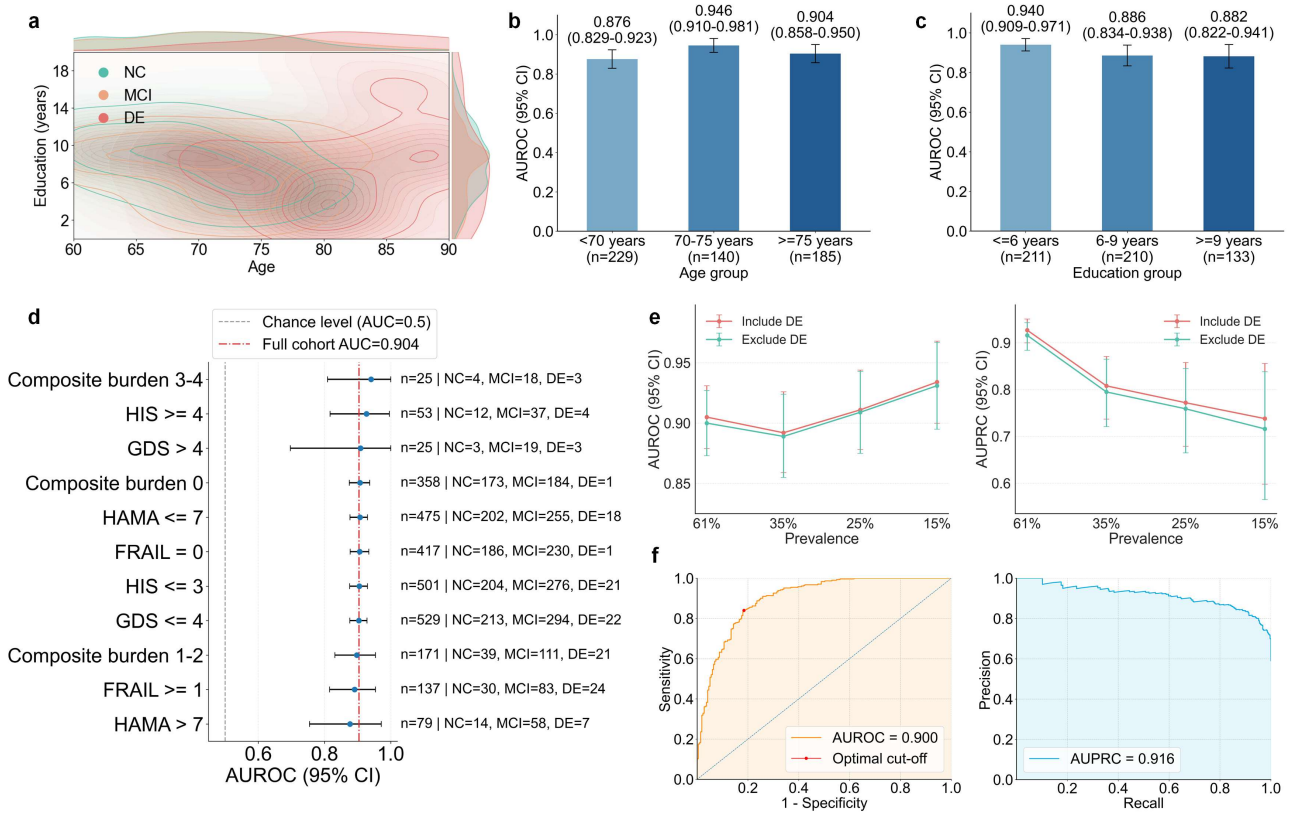


Fig. 4 Robustness of model performance across demographic and clinical subgroups. **a**, Joint distribution of age and years of education across assessment groups in the external validation cohort. **b**, Age-stratified subgroup analysis showing stable model discrimination across < 70, 70-75 and ≥ 75 years, with AUROCs of 0.876, 0.946 and 0.904, respectively. **c**, Education-stratified subgroup analysis across 1-6, 6-9 and ≥ 9 years of education, with AUROCs of 0.940, 0.886 and 0.882, respectively. **d**, Forest plot of AUROC estimates across clinical and health-related subgroups defined by ischemic burden (HIS), anxiety symptoms (HAMA), depressive symptoms (GDS) and frailty status (FRAIL). A composite burden score was further constructed by summing four abnormal domains, including HIS ≥ 4, HAMA > 7, GDS > 4 and FRAIL ≥ 1, and categorized as 0, 1-2 and 3-4 to represent increasing gradients of subclinical burden. **e**, Comparison of AUROC and AUPRC under different prevalence scenarios before and after excluding dementia cases. **f**, NC versus MCI discrimination after excluding dementia cases, showing preserved early-stage classification performance with an AUROC of 0.900 and an AUPRC of 0.916. Error bars and horizontal intervals indicate 95% CIs.

To assess robustness and potential demographic bias prior to clinical deployment, we performed stratified analyses by age, education level, and clinical indicators in the external validation cohort. We first compared the distributions of age and years of education across assessment groups (Fig. 4a). Across

age strata (< 70, 70-75, ≥ 75 years), the model achieved AUROCs of 0.876, 0.946, and 0.904, respectively (Fig. 4b); across education strata (≤ 6, 6-9, ≥ 9 years), AUROCs were 0.940, 0.886, and 0.882 (Fig. 4c), indicating stable performance across varying levels of cognitive reserve. We further stratified by anxiety (HAMA), depressive symptoms (GDS), ischemic burden (HIS), and frailty (FRAIL). The resulting forest plot (Fig. 4d) shows AUROC point estimates clustering closely around the full-cohort value, with overlapping confidence intervals across subgroups, confirming that discriminative ability is preserved across a heterogeneous range of mood, vascular, and frailty profiles.

We additionally restricted the analysis to NC versus MCI classification by excluding all DE individuals. The model retained an AUROC of 0.900 for NC-MCI discrimination (Fig. 4f) with comparable trends across simulated prevalence scenarios (Fig. 4e), demonstrating that outputs do not rely on coarse dementia-driven differences. Taken together, these findings suggest that the model's discriminative power arises primarily from dynamic, task-evoked multimodal physiological and behavioral signals rather than from static demographic factors such as age or education.

2.6 Clinical interpretation through cognitive function quantification

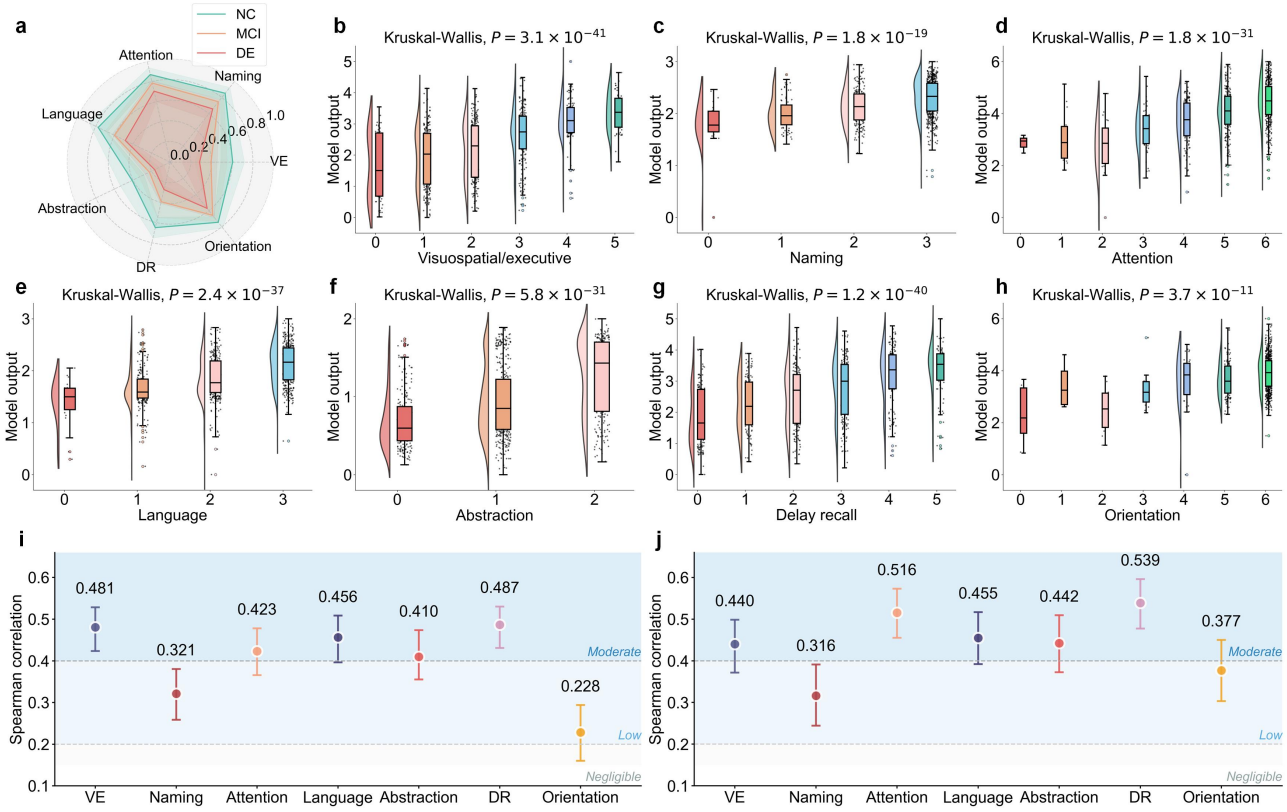


Fig. 5 Alignment of model-derived and MoCA subdomain scores across seven cognitive functions. a, Radar plots of average model-derived scores across seven cognitive functions (visuospatial/executive, naming, attention, language, abstract reasoning, delayed recall, and orientation) in the internal development cohort, showing clear group differences. **b–h**, Distribution of model-derived scores for each domain relative to MoCA subdomain scores in the development cohort, with box and violin plots demonstrating consistent score increases with higher MoCA levels. Kruskal–Wallis tests [42] confirm significant group differences (all $P < 0.001$). **i**, Correlation between model-derived and MoCA scores in the internal development cohort. **j**, Validation results in the external cohort. Correlations represent Spearman's r with 95% CIs derived from bootstrapping ($n = 500$); all $P < 0.001$.

After evaluating model performance, generalizability, and subgroup robustness, we further assessed whether the model outputs were interpretable at the cognitive function level. In addition to estimating overall cognitive impairment status, the model generated scores for seven commonly assessed cognitive functions: visuospatial/executive function (VE), naming, attention, language, abstraction, delayed recall (DR), and orientation [39]. As shown in Fig. 5a, the model output profiles showed a graded decline from NC to MCI and DE, especially in VE, attention, DR, abstraction, and language, consistent with increasing cognitive impairment severity across groups.

Fig. 5b-j further compares the model-derived cognitive function scores with the corresponding MoCA cognitive function scores. Participants with higher MoCA scores for each cognitive function generally had higher model-derived scores, and significant differences were observed across score-level groups by Kruskal-Wallis tests [42] (all $P < 0.001$). These results suggest that the model does not only produce an overall status label, but also provides an interpretable cognitive function profile aligned with established neuropsychological assessment dimensions. Such outputs may help describe which cognitive functions are more affected in each individual and support more fine-grained assessment recommendations.

3 Discussion

Our study demonstrates how specific cognitive tasks evoke distinct cross-modal dynamic patterns. We established a mapping mechanism between these transient physiological-behavioral changes and specific cognitive functions, such as executive function, memory, and attention. By synchronizing EEG, ECG, speech, and video signals during structured tasks, we quantified the "physiological-behavioral response" associated with cognitive impairment [9, 43]. Our results suggest that these multimodal features do not exist in isolation. Instead, they synergistically characterize the cognitive state: neurophysiological signals reflect intrinsic cognitive load and compensatory regulation, while behavioral signals capture explicit performance adaptation. This confirms that analyzing task-evoked multimodal signals provides an objective, interpretable basis for assessing specific cognitive functions.

To elucidate the pathological link between multimodal signals and cognitive impairment, we mapped task-evoked features to specific cognitive functions. The results showed clear modality- and task-dependent patterns. EEG and speech were mainly associated with higher-order cognitive domains, including attention, and language, while EEG/ECG complexity and entropy were closely related to visuospatial/executive and attention (Fig. 2). These findings are consistent with the "physiological loss of complexity" hypothesis [44, 45], recent evidence on wearable EEG for MCI detection [25], and the association between heart rate variability and cognitive impairment [37]. Speech prosody and semantic features were further linked to language and memory domains [35, 46] (Fig. 2), while video-derived facial kinematics and eye-tracking patterns provided complementary information on emotional regulation [47] and abstract reasoning deficits.

Compared with existing digital biomarker studies, our framework differs in two aspects. First, previous passive sensing approaches, including GPS-based driving trajectories [36], continuous actigraphy [33], and long-term multimodal monitoring systems [38], are suitable for longitudinal observation. In contrast, our system adopts an active task-evoked paradigm, using structured cognitive loads to elicit task-specific physiological and behavioral responses within a standardized assessment window. Second, compared with unimodal approaches such as spontaneous speech-based acoustic and interaction analysis [34], our method synchronously integrates EEG, ECG, speech, and video signals during cognitive tasks. This design enables the model to capture task-induced response trajectories, in which high-load tasks such as calculation (T4) and target search (T10) may capture additional task-related information relative to resting-state features [48]. These dynamic changes in EEG spectral power [49], ECG heart rate variability [28, 50], facial kinematics, and eye movements [51] may support a multidimensional characterization of cognitive impairment across neural, autonomic, and behavioral levels.

These findings provide a complementary perspective on the translation of wearable multimodal systems into community and clinical settings. The proposed system is intended to complement existing approaches for cognitive impairment assessment by quantifying neurophysiological dysregulation, reflected by EEG and ECG signals, and variations in behavioral adaptability, reflected by speech and facial features, during structured cognitive task interactions. Conventional neuropsychological tests primarily evaluate the final cognitive performance outcomes, whereas the task-evoked multimodal approach proposed in this study is designed to reflect the underlying physiological compensatory load required to maintain task performance. In particular, under specific task conditions, such as calculation and target search tasks (Table 2 and Supplementary Table 3), the model showed relatively high sensitivity. By quantifying low-level physiological dynamics and millisecond-level behavioral variations that are difficult to capture through routine clinical observation, this system provides complementary information and suggests the potential of task-driven characterization of physiological and behavioral differences associated with cognitive impairment.

Importantly, this model may complement current clinical workflows that rely mainly on conventional screening scales and clinical history. Unlike traditional scales, which primarily depend on explicit cognitive outputs, the multimodal system captures integrated physiological and behavioral signatures during task execution and may help reduce, but not eliminate, the influence of educational background and practice effects [52]. For individuals within the assessment gray zone, such as those with borderline scale scores due to educational barriers or subjective masking effects [53, 54], the system provides an objective physiological-behavioral dimension that may support more objective triage and clinical stratification (Supplementary Fig. 25). In addition to serving as an auxiliary front-end measurement, the system can characterize fine-grained features associated with core cognitive domains during task interaction. In future applications, this capability may support dynamic monitoring of physiological or behavioral changes during targeted non-pharmacological interventions, such as personalized cognitive rehabilitation or audiovisual stimulation.

To better define the practical utility of this multimodal system in a broader public health context, Geoffrey Rose's population strategy [55] was considered. This framework helps contextualize the potential public health relevance of the present work: from a population-risk perspective, future dementia burden may arise substantially from large middle-risk populations rather than only from a small high-risk subgroup. Therefore, a low-burden assessment tool that establishes objective functional baselines in community populations may complement standard cognitive health risk management. Moreover, as demonstrated by Walsh et al. [56], the clinical trajectory of community-based mild cognitive impairment is dynamic and may include reversion. Thus, individuals identified as being at intermediate or high risk by the proposed system should not be interpreted as carrying an irreversible pathological label. Instead, this classification may indicate an opportunity for timely follow-up assessment and individualized health management, providing primary care practitioners with multidimensional functional information to support subsequent decision-making.

The translation of digital biomarkers into public health infrastructure requires careful consideration of their social and ethical implications [57]. The purpose of this system is not to establish a fully automated assessment center with limited interpersonal interaction. Rather, by supporting the time-consuming process of standardized multidomain assessment, the system may help reduce the administrative burden on primary care practitioners and preserve meaningful patient-clinician communication during clinical consultation. Given the particular concerns regarding privacy and adherence among older adults, an active and controlled sensing paradigm was adopted, in which data acquisition was restricted to a brief and supervised time window. Nevertheless, non-response, task non-completion, task abandonment, and insufficient multimodal data quality may themselves represent clinically meaningful informative missingness in older-adult community cohorts. To further address privacy concerns related to multimedia audio and video recordings, structured features were extracted through local deployment, thereby minimizing the need to store or transmit highly sensitive raw audio and video data. However, in real-world deployment, probability-stratification outputs may increase psychological burden, particularly in relation to the right not to know, and may also create risks of discrimination in employment or insurance-related contexts if adequate policy safeguards are absent. Therefore, future clinical integration of this technology should involve collaboration with ethicists and prioritize strict data governance, transparent consent procedures, and equitable accessibility.

On this basis, extending the system from controlled community centers to less controlled real-world environments represents an important future direction [6]. The task-modality-feature associations and specificity analyses identified in this study may inform the design of a context-aware sensing paradigm, in which specific sensors are selectively activated during relevant daily activities. Future iterations of the system may integrate multimodal sensors into unobtrusive wearable devices equipped with context-aware noise reduction, thereby enabling continuous and low-burden cognitive monitoring in the daily lives of older adults.

Despite the encouraging validation results, several limitations warrant further investigation. First, although the use of dual-channel prefrontal EEG maximized wearability and adherence in community settings, it inherently lacks spatial resolution and is insufficient to comprehensively capture brain network dynamics elicited by high cognitive load [25]. This limitation restricts deeper observation of task-evoked neural responses. In addition, the current analytical framework mainly relies on predefined features to ensure model interpretability and therefore may not fully exploit the synergistic information contained in raw time-series data. Future work will continue to expand the dataset and use the

task-modality-feature-cognitive function associations established in this study as prior knowledge for multimodal representation models or large language model-assisted interpretation [58], with the goal of supporting explainable reasoning for cognitive impairment based on physiological and behavioral signals. Second, the cross-sectional design provides only a static representation of cognitive status. Future studies should focus on longitudinal follow-up to examine the relationship between clinical trajectories and signal evolution, thereby better characterizing physiological and behavioral changes associated with cognitive impairment. Second, the cross-sectional design provides only a static representation of cognitive status. Future studies should focus on longitudinal follow-up to examine the relationship between clinical trajectories and signal evolution. Longitudinal real-world studies should also explicitly account for non-response, incomplete testing, task abandonment, and modality-level missingness, as these factors may themselves reflect clinically relevant functional limitations. Third, direct links to neurobiology have not yet been established. Future work should integrate wearable-derived indicators with neuroimaging and fluid biomarkers [18] to explore the neuropathological relevance of these physiological and behavioral signals.

In conclusion, this study establishes a comprehensive cognitive impairment assessment system based on task-evoked multimodal wearable sensing. By integrating EEG, ECG, speech, and video data captured during structured multidimensional cognitive tasks, the system enables fine-grained evaluation across seven core cognitive functions and effective probability stratification. Experimental results on both development and independent external validation cohorts demonstrate the model's robustness across varying prevalence settings. The system quantifies task-evoked physiological stress and latent behavioral responses, such as anxiety, even in community settings with limited specialist supervision. These findings support the potential of wearable multimodal sensing as an auxiliary approach for improving the efficiency and accessibility of cognitive impairment assessment in community-based healthcare.

4 Methods

4.1 Study cohort

This study was conducted from December 2022 to July 2025, with data collection carried out in community settings across Nanjing. A total of 1,608 older adults were recruited from community settings. After data verification and quality control, 1,385 participants were included in the final analysis. Demographic and clinical characteristics of the study population are summarized in Table 1 and Supplementary Table 1. The study protocol was reviewed and approved by the Ethics Committee of Nanjing Hospital of Chinese Medicine (approval number: KY2022004). All procedures adhered to the principles of the Declaration of Helsinki. Prior to assessment, participants were informed in detail about the study objectives, evaluation procedures, potential benefits, and privacy protection measures. Written informed consent was obtained from all participants before any assessments were initiated. For audio and video data, participants and, where applicable, their caregivers were informed of the recording purposes and privacy safeguards. To minimize privacy risks, model development used locally extracted structured features rather than identifiable raw audio or video recordings.

Experimental procedure: Participants provided sociodemographic information, medical history, family history, and lifestyle details (Supplementary Fig. 27). Clinical assessment were adjudicated by experienced neurologists based on comprehensive and standard criteria [5]. Before the multimodal assessment, participants received a comprehensive explanation of the experimental procedures, including an introduction to the data acquisition devices and their functionalities, to help reduce anxiety. During the evaluation, participants were instructed to remain relaxed and comfortable, minimize excessive head or body movements, and avoid generating unrelated sounds. Each participant was seated in front of a tablet and equipped with multimodal sensors to complete the cognitive tasks, as illustrated in Supplementary Fig. 1. The entire process was carried out independently by the participant, with researchers offering only minimal guidance and necessary assistance. The average time required to complete the multimodal assessment with our system was 20.53 ± 2.51 minutes. This total duration encompasses initial preparation and all exploratory tasks (detailed temporal breakdowns for individual tasks are provided in Supplementary Table 12). Notably, our subsequent cross-task analysis revealed that a streamlined protocol utilizing only the five most discriminative tasks can be completed in just 8.34 ± 1.78 minutes while maintaining high classification performance (AUROC = 0.88).

4.2 Wearable multimodal acquisition system

A wearable system was used for the synchronous acquisition of multimodal physiological and behavioral signals [59], including EEG, ECG, speech and video data in Supplementary Fig. 1. The central control unit was a Lenovo TB-J607F tablet, which provided real-time display and recording of all incoming signals. EEG signals were recorded from the prefrontal regions using Fp1 and Fp2 electrodes, positioned according to the international 10–20 system. Signals were captured using dry electrodes at a sampling rate of 250 Hz. For ECG acquisition, a standard three-electrode, single-lead configuration was used, also sampled at 250 Hz. All physiological signal modules supported low-power Bluetooth 5.0 wireless transmission to ensure real-time data synchronization and stability. Speech signals were recorded using a MOMA S1 wireless clip-on microphone at a sampling rate of 48 kHz, supporting dual-channel stereo recording. Video data were captured using the tablet's built-in front-facing camera with a resolution of 1,920 $\times$ 1,080 pixels and a frame rate of 30 frames per second. The system is characterized by its portability and its capability to perform real-time, synchronized multimodal data collection, making it suited for wearable cognitive impairment assessment studies. The multimodal signal acquisition device was evaluated by an independent third-party laboratory to ensure performance consistency and safety compliance in accordance with national software quality standards (GB/T 25000.51-2016).

4.3 Multidimensional cognitive function assessment tasks

To evaluate physiological and behavioral responses in a wearable setting, we developed a tablet-based multidimensional cognitive tasks designed to comprehensively assess various cognitive functions. Further details can be found in Supplementary Table 10, 11. These tasks integrate multimodal inputs, including EEG, ECG, speech, and facial expressions, to capture physiological responses and behavioral characteristics during cognitive task execution. The task design targets seven core cognitive functions: visuospatial/executive, naming, attention, language, abstraction, delayed recall, and orientation as shown in Supplementary Fig. 29. By leveraging simultaneous analysis of multimodal signals, the system enables quantification of individual cognitive states across multiple functional dimensions. This multimodal, multi-domain approach supports comprehensive assessment for cognitive impairment from multiple perspectives.

4.4 Multimodal feature extraction

We established a multimodal signal processing and feature extraction pipeline to analyze synchronously acquired physiological and behavioral data, including EEG, ECG, speech, and video signals. The overall framework is illustrated in Supplementary Fig. 2b. Each modality was independently preprocessed and transformed into task-specific feature sequences, which were subsequently used for cognitive impairment modeling. For EEG, raw signals were band-pass filtered at 0.5-50 Hz, quality-controlled using a sliding-window strategy, corrected for artifacts using wavelet decomposition with quantile normalization, and resampled to 100 Hz [60]. Frequency-domain, nonlinear, time-domain, and waveform morphology features were extracted to characterize cortical activation, rhythmic slowing, neural complexity, signal variability, and transient neurophysiological fluctuations across cognitive tasks (Supplementary Table 14). For ECG, signals were denoised using median filtering and detrending, followed by band-pass and notch filtering [61]. R-peak detection was then performed to derive RR interval sequences, from which cardiac waveform parameters and heart rate variability features were extracted to characterize autonomic regulation and task-evoked cardiac responses (Supplementary Table 15).

For speech, audio signals were resampled to 16 kHz and processed to extract both acoustic and semantic features. Acoustic descriptors were derived from the eGeMAPSv02 feature set [62], covering prosodic, spectral, voice quality, and pause-related characteristics, while semantic features were obtained from automatic transcription and language modeling to quantify contextual embeddings, semantic coverage, sentence similarity, and syntactic depth [63] (Supplementary Table 16). For video, recordings were standardized in format and resolution, followed by face detection, landmark localization, alignment, head pose estimation, and eye-movement tracking. Facial action units, expression entropy, head movement dynamics, gaze trajectory, and blink-related features were extracted to characterize emotional expression, psychomotor behavior, visual attention, and behavioral responsiveness during cognitive tasks [64] (Supplementary Table 17). Together, these modality-specific features provided complementary physiological and behavioral representations for downstream cognitive impairment assessment.

4.5 Multimodal cognitive impairment assessment model

We developed a multimodal assessment model that integrates physiological and behavioral signals to concurrently classify cognitive impairment and estimate multidimensional cognitive function scores. The overall architecture is illustrated in Supplementary Fig. 2a. The model takes multimodal data from 11 structured tasks as input, including EEG, ECG, video sequences, speech features, task metadata, and demographic information. Each modality is first encoded through a modality-specific encoder. For EEG, ECG, and video data, a dynamic task adapter (DTA) module was used to capture task-dependent response patterns across cognitive tasks (Supplementary Fig. 2c), whereas speech, task performance, and demographic features were transformed into unified vector representations through fully connected layers. The resulting modality-specific representations were then integrated using learnable modality weights to generate a joint representation of cognitive status. To improve robustness in real-world settings with incomplete data, the model incorporated a synergistic missing-modality strategy combining structural imputation and adaptive routing (Supplementary Table 19).

To enhance interpretability and align model learning with clinical cognitive domains, we introduced an intermediate seven-dimensional cognitive function supervision module. The seven dimensions corresponded to MoCA-aligned cognitive functions, including visuospatial/executive function, naming, attention, language, abstraction, delayed recall, and orientation. During training, the fused multimodal representation was supervised using the binary labels of these seven cognitive subdomains through a multi-label learning strategy [65]. This design encouraged the model to learn structured patterns associated with cognitive impairment rather than relying only on a single global classification label. During inference, the model produced both an overall cognitive impairment classification and a seven-dimensional cognitive function profile, supporting a more interpretable assessment of cognitive status.

References

[1] Livingston, G. *et al.* Dementia prevention, intervention, and care: 2024 report of the lancet standing commission. *Lancet* **404**, 572–628 (2024).

[2] Jia, L. *et al.* Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in china: A cross-sectional study. *Lancet Public Health* **5**, e661–e671 (2020).

[3] Cornblath, E. J. *et al.* Defining and predicting transdiagnostic categories of neurodegenerative disease. *Nat. Biomed. Eng.* **4**, 787–800 (2020).

[4] Nichols, E. *et al.* Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the global burden of disease study 2019. *Lancet Public Health* **7**, e105–e125 (2022).

[5] Jia, L. *et al.* Dementia in china: Epidemiology, clinical management, and research advances. *Lancet Neurol.* **19**, 81–92 (2020).

[6] Butler, P. M. *et al.* Smartwatch-and smartphone-based remote assessment of brain health and detection of mild cognitive impairment. *Nat. Med.* **31**, 829–839 (2025).

[7] Rafii, M. S. & Aisen, P. S. Detection and treatment of Alzheimer's disease in its preclinical stage. *Nat. Aging* **3**, 520–531 (2023).

[8] Li, S. *et al.* A community-codesigned llm-powered chatbot for primary care: a randomized controlled trial. *Nat. Health* **1**, 238–250 (2026).

[9] Bartrés-Faz, D. *et al.* Psychological profiles associated with mental, cognitive and brain health in middle-aged and older adults. *Nat. Ment. Health* **3**, 92–103 (2025).

[10] Bloomberg, M., Muniz-Terrera, G., Brocklebank, L. & Steptoe, A. Healthy lifestyle and cognitive decline in middle-aged and older adults residing in 14 European countries. *Nat. Commun.* **15**, 5003 (2024).

[11] Chan, A. T., Ip, R. T., Tran, J. Y., Chan, J. Y. & Tsoi, K. K. Computerized cognitive training for memory functions in mild cognitive impairment or dementia: A systematic review and meta-analysis. *npj Digit. Med.* **7**, 1 (2024).

[12] Liew, T. M. *et al.* PENSIEVE-AI a brief cognitive test to detect cognitive impairment across diverse literacy. *Nat. Commun.* **16**, 2847 (2025).

[13] Lerede, A. *et al.* Large-scale online assessment uncovers a distinct Multiple Sclerosis subtype with selective cognitive impairment. *Nat. Commun.* **16**, 6938 (2025).

[14] Zhang, J. *et al.* Five-minute cognitive test as a new quick screening of cognitive impairment in the elderly. *Aging Dis.* **10**, 1258 (2019).

[15] Chen, S. *et al.* Automatic dementia screening and scoring by applying deep learning on clock-drawing tests. *Sci. Rep.* **10**, 20854 (2020).

[16] Xue, C. *et al.* AI-based differential diagnosis of dementia etiologies on multimodal data. *Nat. Med.* **30**, 2977–2989 (2024).

[17] Oh, H. S.-H. *et al.* A cerebrospinal fluid synaptic protein biomarker for prediction of cognitive resilience versus decline in Alzheimer's disease. *Nat. Med.* **31**, 1592–1603 (2025).

[18] Valletta, M. *et al.* Blood biomarkers of Alzheimer's disease and progression across different stages of cognitive decline in the community. *Nat. Commun.* **16**, 10412 (2025).

[19] Grande, G. *et al.* Blood-based biomarkers of Alzheimer's disease and incident dementia in the community. *Nat. Med.* **31**, 2027–2035 (2025).

[20] Langella, S. *et al.* Effect of apolipoprotein genotype and educational attainment on cognitive function in autosomal dominant Alzheimer's disease. *Nat. Commun.* **14**, 5120 (2023).

[21] Andrieu, S. *et al.* Harnessing artificial intelligence to transform Alzheimer's disease research. *Nat. Med.* **31**, 1384–1385 (2025).

[22] Lu, Y. *et al.* Wearable data link urban green space to physical activity. *Nat. Health* **1**, 67–77 (2025).

[23] Qi, W. *et al.* Alzheimer's disease digital biomarkers multidimensional landscape and AI model scoping review. *npj Digit. Med.* **8**, 366 (2025).

[24] Berron, D. *et al.* A remote digital memory composite to detect cognitive impairment in memory clinic samples in unsupervised settings using mobile devices. *npj Digit. Med.* **7**, 79 (2024).

[25] He, C., Yu, X., Zhang, Y., Li, Y. & Jiang, N. Wearable eeg devices in the detection of mild cognitive impairment: a systematic review. *npj Digit. Med.* **9**, 265 (2026).

[26] Anjum, M. F. *et al.* Resting-state EEG measures cognitive impairment in Parkinson's disease. *npj Parkinsons Dis.* **10**, 6 (2024).

[27] Tan, E. S. *et al.* Association of atrial electrophysiological abnormalities with cognitive decline and cerebrovascular disease. *J. Am. Heart Assoc.* **15**, e045054 (2026).

[28] Arco, J. E., Gallego-Molina, N. J., Ortiz, A., Arroyo-Alvis, K. & López-Pérez, P. J. Identifying HRV patterns in ECG signals as early markers of dementia. *Expert Syst. Appl.* **243**, 122934 (2024).

[29] Azadmaleki, H. *et al.* SpeechCARE: dynamic multimodal modeling for cognitive screening in diverse linguistic and speech task contexts. *npj Digit. Med.* **8**, 677 (2025).

[30] Shankar, R., Goh, Z., Devi, F. & Xu, Q. A systematic review of explainable artificial intelligence methods for speech-based cognitive decline detection. *npj Digit. Med.* **8**, 724 (2025).

[31] Gerbaudo-González, N., Condorí-Fernández, N., Catalá, A. & Gandoy-Crego, M. AI-based emotion recognition in dementia through facial expression: A scoping review. *Ageing Res. Rev.* **114**, 102975 (2026). Epub 2025 Nov 27.

[32] Desai, R. *et al.* Affective symptoms and risk of progression to mild cognitive impairment or dementia in subjective cognitive decline: A systematic review and meta-analysis. *Ageing Res. Rev.* **71**, 101419 (2021).

[33] Li, J. *et al.* TATC: Predicting Alzheimer's disease with actigraphy data. Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (2018).

[34] Nasreen, S., Hough, J. & Purver, M. Detecting Alzheimer's disease using interactional and acoustic features from spontaneous speech. Proc. Interspeech 2021 (2021).

[35] Jafari, Z., Andrew, M. K. & Rockwood, K. J. Diagnostic utility of speech-based biomarkers in mild cognitive impairment: A systematic review and meta-analysis. *Age Ageing* **54**, afaf316 (2025).

[36] Bayat, S. *et al.* GPS driving: A digital biomarker for preclinical Alzheimer disease. *Alzheimers Res. Ther.* **13**, 115 (2021).

[37] Derby, C. A. *et al.* Heart rate variability is associated with global cognition and cognitive impairment: Results from the einstein aging study. *Alzheimers Dement.* **19**, e064226 (2023).

[38] Ouyang, X. *et al.* ADMarker: A multi-modal federated learning system for monitoring digital biomarkers of Alzheimer's disease. Proc. 30th Annu. Int. Conf. Mob. Comput. Netw. (2024).

[39] Nasreddine, Z. S. *et al.* The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53**, 695–699 (2005).

[40] Arevalo-Rodriguez, I. *et al.* Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst. Rev.* **7**, CD010783 (2015).

[41] DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **44**, 837–845 (1988).

[42] Yang, Y. *et al.* Artificial intelligence-enabled detection and assessment of Parkinson's disease using nocturnal breathing signals. *Nat. Med.* **28**, 2207–2215 (2022).

[43] Zhang, Y. *et al.* Identification of psychiatric disorder subtypes from functional connectivity patterns in resting-state electroencephalography. *Nat. Biomed. Eng.* **5**, 309–323 (2021).

[44] Deng, J. *et al.* Novel methodology for detection and prediction of mild cognitive impairment using resting-state EEG. *Alzheimers Dement.* **20**, 145–158 (2024).

[45] Jiao, B. *et al.* Neural biomarker diagnosis and prediction to mild cognitive impairment and Alzheimer's disease using EEG technology. *Alzheimers Res. Ther.* **15**, 32 (2023).

[46] Yang, Q., Li, X., Ding, X., Xu, F. & Ling, Z. Deep learning-based speech analysis for Alzheimer's disease detection: A literature review. *Alzheimers Res. Ther.* **14**, 186 (2022).

[47] Fei, Z. *et al.* A novel deep neural network-based emotion analysis system for automatic detection of mild cognitive impairment in the elderly. *Neurocomputing* **468**, 306–316 (2022).

[48] Montero-Odasso, M. *et al.* Quantitative gait analysis under dual-task in older people with mild cognitive impairment: A reliability study. *J. Neuroeng. Rehabil.* **6**, 1–6 (2009).

[49] Azami, H. *et al.* Beta to theta power ratio in EEG periodic components as a potential biomarker in mild cognitive impairment and Alzheimer's dementia. *Alzheimers Res. Ther.* **15**, 133 (2023).

[50] Mao, M. *et al.* Ventricular electrocardiographic signatures associated with dementia and plasma Alzheimer's biomarkers in older adults: The MIND-China study. *Alzheimers Dement.* **19**, e060327 (2023).

[51] Xu, X. *et al.* Facial aging, cognitive impairment, and dementia risk. *Alzheimers Res. Ther.* **16**, 245 (2024).

[52] Franzen, S. *et al.* A systematic review of neuropsychological tests for the assessment of dementia in non-western, low-educated or illiterate populations. *J. Int. Neuropsychol. Soc.* **26**, 331–351 (2020).

[53] Elman, J. A. *et al.* Practice effects persist over two decades of cognitive testing: Implications for longitudinal research. *Alzheimers Dement.* **21**, e106014 (2025).

[54] Ko, E., Kim, S. & Wu, B. Help-seeking behaviors in dementia caregivers of east asian backgrounds: An integrative review. *Alzheimers Dement.* **21**, e097776 (2025).

[55] Frank, J. W. Controlling the obesity pandemic: Geoffrey Rose revisited. *Can. J. Public Health* **113**, 736–742 (2022).

[56] Walsh, S. *et al.* Population-level interventions for the primary prevention of dementia: A complex evidence review. *Lancet* **402**, S13–S14 (2023).

[57] Ford, E., Milne, R. & Curlewis, K. Ethical issues when using digital biomarkers and artificial intelligence for the early detection of dementia. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **13**, e1492 (2023).

[58] Rutunda, S. *et al.* Large language models for frontline healthcare support in low-resource settings. *Nat. Health* **1**, 191–197 (2026).

[59] Zhao, M. *et al.* Multi-query cross-modal attention fusion for cognitive impairment recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **33**, 2520–2530 (2025).

[60] Song, T., Zheng, W., Song, P. & Cui, Z. EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans. Affect. Comput.* **11**, 532–541 (2018).

[61] Sarkar, P. & Etemad, A. Self-supervised ECG representation learning for emotion recognition. *IEEE Trans. Affect. Comput.* **13**, 1541–1554 (2020).

[62] Eyben, F. *et al.* The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Trans. Affect. Comput.* **7**, 190–202 (2015).

[63] Fard, A. P., Mahoor, M. H., Alsuhaibani, M. & Dodge, H. H. Linguistic-based mild cognitive impairment detection using informative loss. *Comput. Biol. Med.* **176**, 108606 (2024).

[64] Tian, Y., Kanade, T. & Cohn, J. Recognizing action units for facial expression analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 97–115 (2001).

[65] Liu, W., Wang, H., Shen, X. & Tsang, I. W. The emerging trends of multi-label learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 7955–7974 (2021).

5 Declarations

Data availability

The datasets generated and analyzed during this study are currently being organized and reviewed for compliance with ethical and privacy regulations. Upon completion of the review process, de-identified data may be made publicly available. In the meantime, data access can be requested from the corresponding author (chengyu@seu.edu.cn) with a formal data use proposal. Once approved, access will be granted under a data use agreement. A minimal de-identified dataset sufficient to reproduce the key results is available in the Code Ocean capsule for the peer review process.

Code availability

All custom scripts and implementation details used in this study are available at: https://github.com/zmh56/M3-CIA-Feature-Engineering. The Code Ocean capsule provided through the submission system includes all code, dependencies, and sample data required for one-click execution and verification of the results during peer review.

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China (Grant No. 2023YFC3603600), and in part by the National Natural Science Foundation of China (Grant Nos. 62171123). We thank our collaborators at Nanjing Hospital of Chinese Medicine for their valuable assistance in data collection and project coordination.

Author contributions

C.L. and J.L. conceived and supervised the study, provided overall conceptual guidance, and oversaw manuscript preparation. Y.B. and T.Z. provided guidance on clinical and model development. H.Y. and Y.S. were in charge of data collection. C.L. and M.Z. designed the experiments and contributed to manuscript writing. M.Z., L.Z., C.Y., Y.F., J.Z. and H.G. processed multimodal signals and performed clinical validation. L.Z., L.P. and W.X. contributed to the experimental foundation, device development, and manuscript revision. Y.L., Y.X. and H.X. contributed to the manuscript revision. H.Y., X.Q. and J.G. designed the evaluation protocols and conducted the clinical experiments. Y.Z., S.T. and H.L. provided clinical expertise and medical oversight, including guidance on assessment standards. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Ethics declarations

The study involving human participants was approved by the Ethics Committee of Nanjing Hospital of Chinese Medicine (Approval No. KY2022004). All participants provided written informed consent in accordance with the Declaration of Helsinki.

Additional information

Supplementary material, including additional figures, tables, and methods, is included as a separate PDF. Corresponding ethical approval documents, device specifications, and software testing reports are also submitted as supplementary materials.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.pdf
- WearableMultimodalCognitiveImpairmentAssessmentSystem.mp4
- ComplianceDocuments.pdf