ARTICLE

Clinical trajectories preceding incident dementia up to 15 years before diagnosis: a large prospective cohort study

Jia You[1,7], Yu Guo[1,7], Yu-Jia Wang[1], Yi Zhang[1], Hui-Fu Wang[1], Lin-Bo Wang[1], Ju-Jiao Kang[1], Jian-Feng Feng[1,2,3,4,5]✉, Jin-Tai Yu[1]✉ and Wei Cheng[1,2,5,6]✉

BACKGROUND: Dementia has a long prodromal stage with various pathophysiological manifestations; however, the progression of pre-diagnostic changes remains unclear. We aimed to determine the evolutional trajectories of multiple-domain clinical assessments and health conditions up to 15 years before the diagnosis of dementia.
METHODS: Data was extracted from the UK-Biobank, a longitudinal cohort that recruited over 500,000 participants from March 2006 to October 2010. Each demented subject was matched with 10 healthy controls. We performed logistic regressions on 400 predictors covering a comprehensive range of clinical assessments or health conditions. Their evolutional trajectories were quantified using adjusted odds ratios (ORs) and FDR-corrected *p*-values under consecutive timeframes preceding the diagnosis of dementia.
FINDINGS: During a median follow-up of 13.7 [Interquartile range, IQR 12.9–14.2] years until July 2022, 7620 subjects were diagnosed with dementia. In general, upon approaching the diagnosis, demented subjects witnessed worse functional assessments and a higher prevalence of health conditions. Associations up to 15 years preceding the diagnosis comprised declined physical strength (hand grip strength, OR 0.65 [0.63–0.67]), lung dysfunction (peak expiratory flow, OR 0.78 [0.76–0.81]) and kidney dysfunction (cystatin C, OR 1.13 [1.11–1.16]), comorbidities of coronary heart disease (OR 1.78 [1.67–1.91]), stroke (OR 2.34 [2.1–1.37]), diabetes (OR 2.03 [1.89–2.18]) and a series of mental disorders. Cognitive functions in multiple tests also demonstrate decline over a decade before the diagnosis. Inadequate activity (3–5 year, overall time of activity, OR 0.82 [0.73–0.92]), drowsiness (3–5 year, sleep duration, OR 1.13 [1.04–1.24]) and weight loss (0–5 year, weight, OR 0.9 [0.83–0.98]) only exhibited associations within five years before the diagnosis. In addition, serum biomarkers of enriched endocrine, dysregulations of ketones, deficiency of brand-chain amino acids and polyunsaturated fatty acids were found in a similar prodromal time window and can be witnessed as the last pre-symptomatic conditions before the diagnosis.
INTERPRETATION: Our findings present a comprehensive temporal-diagnostic landscape preceding incident dementia, which could improve selection for preventive and early disease-modifying treatment trials.

Molecular Psychiatry; https://doi.org/10.1038/s41380-024-02570-0

INTRODUCTION

Dementia ranks seventh among the top ten global causes of death, according to the World Health Organization's 2020 report. Over 55 million people worldwide suffer from dementia, and this staggering number will continue to rise as the growth of aging population and is expected to reach 152 million by 2050 [1]. Dementia imposes tremendous health and economic burden on individuals, their families, and wider society, and has become a global public health priority [2]. The neurodegenerative diseases are progressive with a latency period of decades prior to the diagnosis [3]. The long asymptomatic period, which provides a substantial time window to facilitate access to disease prevention and treatment, has gained growing attention, especially in the current context where dementia is incurable.

Immense efforts have been devoted to detecting the pathophysiological processes underlying dementia that precede the onset of clinical symptoms [1, 2, 4]. However, the early identification and labeling of biomarker-positive patients ahead of symptom occurrence remains questionable due to the uncertainty of progression to dementia [5]. Towards addressing this issue, recent studies have anchored the pre-diagnostic period and unveiled the associations of cognitive and functional impairment [6], depressive symptoms [7], and comorbid conditions [3, 8] with incident dementia. Yet, many other early features

[1]Institute of Science and Technology for Brain-Inspired Intelligence, Department of Neurology, Huashan Hospital, State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Fudan University, Shanghai, China. [2]Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, Shanghai, China. [3]Zhangjiang Fudan International Innovation Center, Shanghai, China. [4]School of Data Science, Fudan University, Shanghai, China. [5]Fudan ISTBI—ZJNU Algorithm Centre for Brain-inspired Intelligence, Zhejiang Normal University, Zhejiang, China. [6]Shanghai Medical College and Zhongshan Hospital Immunotherapy Technology Transfer Center, Shanghai, China. [7]These authors contributed equally: Jia You, Yu Guo. ✉email: jianfeng64@gmail.com; jintai_yu@fudan.edu.cn; wcheng.fdu@gmail.com

Received: 23 May 2023 Revised: 16 April 2024 Accepted: 17 April 2024
Published online: 27 April 2024

and risk factors have been overlooked. Pre-diagnostic changes in daily activities, behaviors, psychological factors, and hematologic measures remain poorly understood. These previous works have narrowly focused on only single symptoms or certain domains of dementia, while investigations designed to capture early systemic signs and to understand the relative orderings of changes across different domains are lacking. In addition, most existing studies often paid attention to the relation of baseline features or factors to dementia, whereas their dynamic progressions on specific subtypes of dementia have been less studied. Although population-based observational studies have illuminated many factors that increase the risk of dementia [1, 2, 9, 10], it's unclear if these features can be captured before symptoms appear and how long before a diagnosis of dementia they are detectable. Currently, well-characterized cohorts based on the general population with long-term follow-up and extensive phenotypic examinations are scarce as they are typically hard to collect.

The UK Biobank cohort provides a unique platform to characterize the pre-dementia phase with rich phenotypic health-related assessments performed years before diagnosis. Herein, we sought to illustrate the evolutionary trajectories of multi-domain features preceding up to 15 years before the first diagnosis of dementia. Further sensitive analysis for Alzheimer's disease (AD) and vascular dementia (VD) was performed, as well. Through this extensive survey, we attempt to validate some established risk factors and early symptomatic manifestations of dementia and to showcase novel candidates for consideration in future anti-dementia strategies.

METHODS

Study design and participants

This study extracted data from the UK-Biobank, a longitudinal cohort with over 500,000 individuals aged between 40 and 69 years at their baseline assessment from March 2006 to October 2010. Participants in the cohort were enrolled from 22 recruitment centers across the UK and underwent multiple examinations. As shown in Fig. 1, participants were excluded if they were: (1) without follow-up records ($n = 1298$); (2) with pre-existing dementia ($n = 250$); (3) diagnosed with dementia over 15 years from their baseline visits ($n = 14$). We performed a matched case-control analysis by identifying ten healthy controls for each demented subject. The matching scheme leveraged rigid criteria of age (± 2 years), gender, and longer follow-up time, and any demented participants cannot match 10 controls under rigid criteria were excluded ($n = 132$); in addition, we adopted several non-rigid criteria, including ethnicity, years of education (± 1 year), and socioeconomic status (Townsend deprivation index $\pm$ 0.5). This study finally included 7620 subjects diagnosed with dementia within 15 years after their baseline visits and 76,200 healthy controls. The UK Biobank has research tissue bank approval from the North West Multi-center Research Ethics Committee and provided oversight for this study. Written informed consent was obtained from all participants.

Target outcomes

The primary outcome was dementia, and its two common subtypes, Alzheimer's disease (AD) and vascular dementia (VD), were studied as the secondary outcomes. The last recorded date was July 22nd, 2022. The outcomes were ascertained and classified according to the ICD-10 codes (Supplementary eTable 1), extracted from UKB clinical records using algorithmically defined data (fields 42018-25), death register documentation (fields 40001-40002), and first occurrences data (fields 131036-37, 130836-43) that compiled from data sources of hospital inpatient, primary care, death certificate and self-report. The date of any earliest recorded diagnosis listed above was leveraged as a proxy for the actual date of diagnosis. Abbreviation, ACD all-cause dementia; AD Alzheimer's dementia; IQR interquartile range; VD vascular dementia.

Pre-diagnosis clinical assessments and health conditions

This study included all clinically relevant variables after excluding those whose missing values were greater than 85%. A total of 400 assessments categorized into eight groups were adopted: activity & sleep ($n = 26$), cognitive functions ($n = 11$), mental health ($n = 20$), physical functions ($n = 84$), self-reported health conditions ($n = 18$), metabolomic blood tests ($n = 168$), serum tests ($n = 63$), and comorbid conditions ($n = 10$). Please refer to eMthods and Supplementary eTable 2 for detailed test procedures of measurements.

Longitudinal data interpretation

Follow-up visits commenced from the date of attendance at the assessment center (field 53), the time that a comprehensive list of clinical assessments were collected. Longitudinal recorded diagnosis of dementia, AD and VD was last updated until July 22nd, 2022. To interpret the trajectories of clinical assessments, we reformatted the data using a nested case-control study strategy. As illustrated in Fig. 2, the analysis was designed by initially aligning the index date of all dementia subjects. Then they can be stratified into different groups of timeframes preceding the diagnosis date. For healthy controls, we assumed they had the same durations from baseline to arbitrary "dementia index dates" as that of the matched dementia subject. Thus, we had multiple groups of paired case-control populations under consecutive time windows preceding the index date of dementia, allowing further investigations at different timeframes.

Statistical analysis

We hypothesized that individuals have relatively worse functional measurements and more health conditions when getting closer to their diagnosis date. Thus, to determine the independent contributions of each predictor to the diagnosis of dementia, we leveraged the population within a five-year time window before their diagnosis. We performed multivariate logistic regressions with adjusted covariates of age at the baseline visit, gender, ethnicity, years of education, Townsend deprivation index, and baseline to index years. *P* values were corrected for multiple comparisons using the false discovery rate (FDR). Predictors with FDR-corrected $p < 0.05$ were adopted for further analysis.

To find trajectories of different assessments before the diagnosis, we considered ten timeframes preceding the dementia index date (Years 0–3, 3–5, 5–6, 6–7, …, 11–12, 12–15); specifically, timeframes of 0–3, 3–5 and 12–15 years were grouped due to insufficient numbers of dementia subjects. We repeated the regression analyses under each timeframe. Thus, for each assessment, we obtained a series of odds ratios with corresponding two-sided *p*-values and tried to determine the time window of significance by using a backward screening strategy on the FDR-corrected *p*-values starting from 3–5 years to 12–15 years before the dementia index. Such a scheme stopped if two of three consecutive timeframes became insignificant (FDR-corrected $p > 0.05$), and we then took the stopping timeframe as the time point of appearing significance between dementia subjects and controls. Noted, all predictors have surpassed the selection criteria under the 0–5 years' timeframe; thus, for any case that happened not significant under 0–3 years, we reported the statistics of 0–5 years.

To have an overview of trajectories of health conditions preceding diagnosis of dementia, Alzheimer's disease, and vascular dementia, stacked additional incidence plots were drawn. The additional incidence was defined as the incidence rates of case group over healthy controls. Health conditions were selected from representative assessments or comorbid or self-reported conditions that were significantly associated with target outcomes. Specifically, clinical assessments of continuous variables were binarized to calculate the incidence. This was performed by setting either 0.95 quantiles (assessments with odds ratios>1) or 0.05 quantiles (assessments with odds ratios <1) from the healthy control group as the cut-off of abnormality. The cut-off was sex-based that derived from male and female populations separately. All health conditions reported in the results were quoted with the derived clinical assessments along with its corresponding odds ratios and 95% confidence intervals. See detailed matching in Supplementary eTables 17, 32 and 45 for dementia, AD and VD, respectively.

Ethnicity was grouped as White, Asian, Black, and others (e.g., mixed or other ethnic groups), where individuals were coded as others if their ethnicity information was missing. We used simple imputation of mode to impute the Townsend deprivation index and years of education whose missingness was less than 3%. Individuals with missing values within the variable of interest were dropped. Odds ratios (OR) reported used normalized statistics if they were continuous variables. Individuals were removed if the predictor of interest was missing during association analysis. We report 0–15 year OR for assessments that are significant over all preceding timeframes, and for other predictors we report the OR of the timeframe starting to become significant.

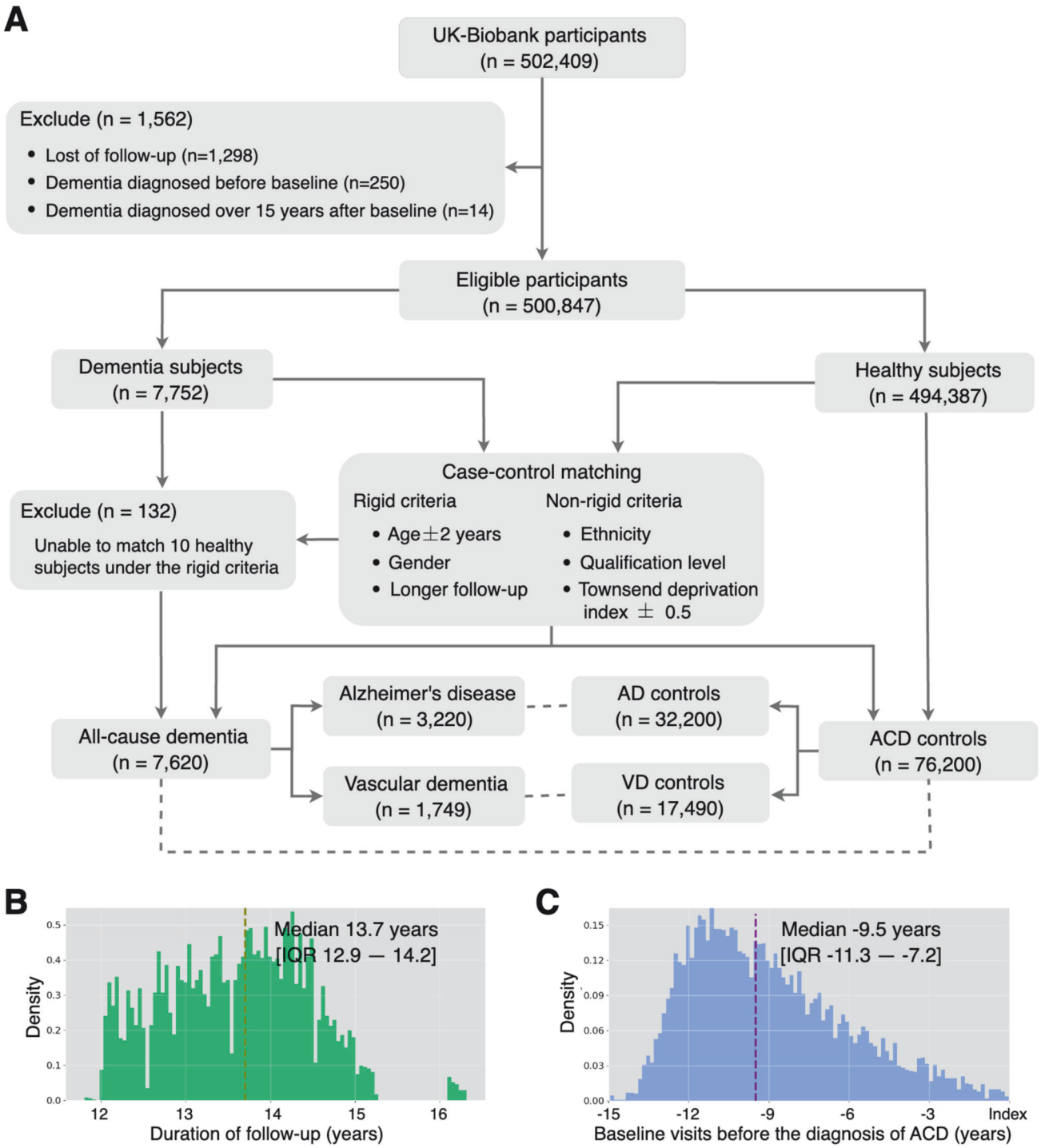

Fig. 1 Flowchart of the case-control participants' selection. A Flowchart of the case-control participants' selection. **B** Histogram of follow-up years (baseline visits to last observation date was July 22nd, 2022). **C** Histogram of the duration of years from baseline visits to the date of dementia index. ACD all-cause dementia, AD Alzheimer's disease, IQR Interquartile range, VD vascular dementia.

RESULTS

Population characteristics

During a median follow-up of 13.7 [Interquartile range, IQR 12.9–14.2] years, we identified 7620 subjects diagnosed with dementia, of whom the diagnosed years were at a median of 9.5 [IQR 7.2–11.3] years after their baseline visits. Among these demented participants, 52.0% were males, and they were mainly comprised of white ethnicity (95.3%) with a mean age of 64.2 [95% CI 64.1–64.3] years at baseline visits. Among the demented individuals, 3220 and 1749 were diagnosed with AD and VD, respectively. Comparable statistics were observed with healthy controls since they were selected based on matched demographic information (Table 1).

Pre-diagnostic clinical assessments and health conditions in dementia

Among 400 available assessments and health conditions, 119 predictors have surpassed the FDR correction within a five-year

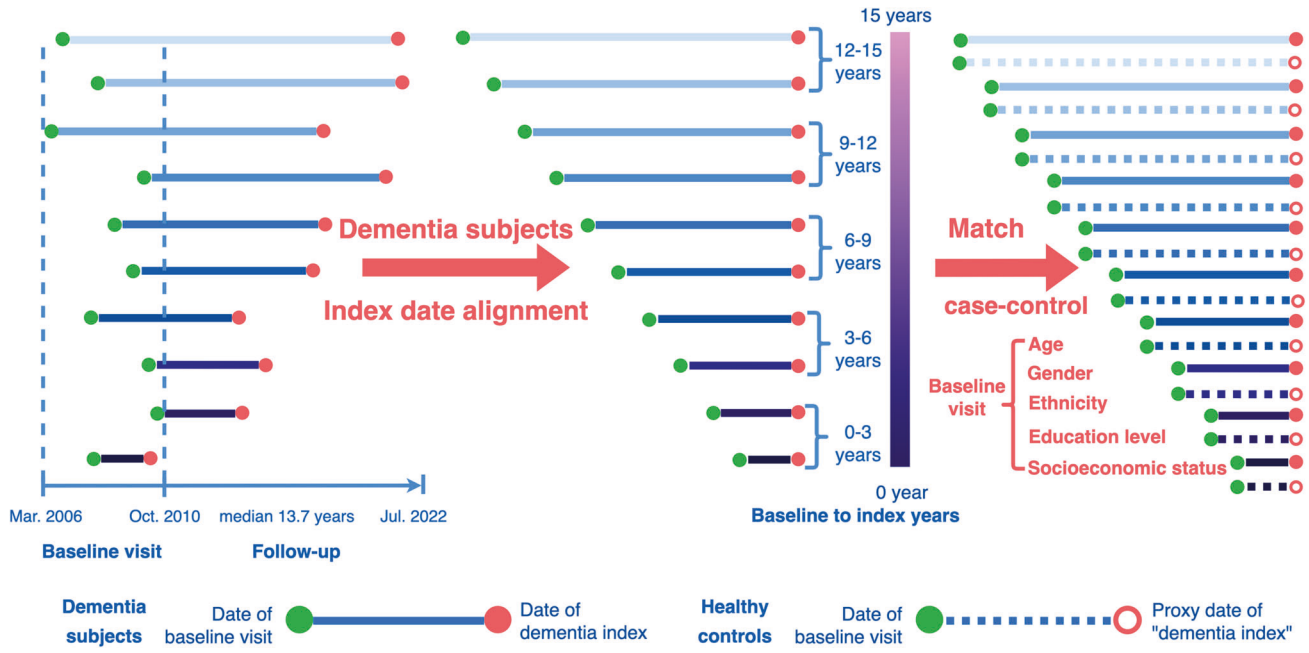

Fig. 2 Study design illustration. Solid horizontal bars represent proxy observational processes of dementia subjects from their baseline visits (green dots) to the dates of dementia index (red dots). We aligned their dementia index dates and stratified the participants based on the duration between their baseline visits to the index dates, from 0 to 15 years (coded with gradient color from dark blue to light blue). A case-control matching was performed under multiple criteria based on the baseline demographic information. For healthy controls, we set an arbitrary date as the proxy date of the "dementia index" (red circles), allowing an equal duration from baseline to "index" to their matched demented subjects.

Table 1. The characteristics of UK-Biobank participants with dementia and healthy controls.

	All-cause dementia		Alzheimer's disease		Vascular dementia	
	demented subjects ($n = 7620$)	Healthy controls ($n = 76,200$)	AD subjects ($n = 3220$)	Healthy controls ($n = 32,200$)	VD subjects ($n = 1749$)	Healthy controls ($n = 17,490$)
Age at baseline (years)	64.2 [64.1, 64.3]	63.4 [63.4, 63.4]	64.7 [64.6, 64.8]	63.8 [63.8, 63.8]	64.8 [64.6, 65.0]	64.0 [63.9, 64.1]
Gender (male, %)	3960 [52.0%]	39,600 [52.0%]	1534 [47.6%]	15,340 [47.6%]	1010 [57.7%]	10,100 [57.7%]
Townsend deprivation index	−0.9 [−1.0, −0.8]	−0.9 [−0.9, −0.9]	−1.1 [−1.2, −1.0]	−1.1 [−1.1, −1.1]	−0.7 [−0.9, −0.5]	−0.8 [−0.8, −0.8]
Ethnicity						
White	7260 [95.3%]	73,007 [95.8%]	3087 [95.9%]	31,026 [96.4%]	1672 [95.6%]	16,808 [96.1%]
Asian	119 [1.6%]	1115 [1.5%]	45 [1.4%]	434 [1.3%]	26 [1.5%]	247 [1.4%]
Black	118 [1.5%]	959 [1.3%]	45 [1.4%]	364 [1.1%]	30 [1.7%]	231 [1.3%]
Others	123 [1.6%]	1119 [1.5%]	43 [1.3%]	376 [1.2%]	21 [1.2%]	204 [1.2%]
Education (years)						
<10	2607 [34.2%]	23,373 [30.7%]	1121 [34.8%]	10,141 [31.5%]	669 [38.3%]	5883 [33.6%]
10–15	2666 [35.0%]	28,959 [38.0%]	1125 [34.9%]	12,172 [37.8%]	590 [33.7%]	6626 [37.9%]
>15	2347 [30.8%]	23,868 [31.3%]	974 [30.2%]	9887 [30.7%]	490 [28.0%]	4981 [28.5%]
Number of individuals (years preceding diagnosis)						
0–3	326 [4.3%]	3260 [4.3%]	79 [2.5%]	790 [2.5%]	76 [4.3%]	760 [4.3%]
3–6	953 [12.5%]	9530 [12.5%]	371 [11.5%]	3710 [11.5%]	223 [13.3%]	2230 [13.3%]
6–9	2062 [27.1%]	20,620 [27.1%]	866 [26.9%]	8660 [26.9%]	445 [25.4%]	4450 [25.4%]
9–12	3172 [41.6%]	31,720 [41.6%]	1412 [43.9%]	14,120 [43.9%]	729 [41.7%]	7290 [41.7%]
12–15	1107 [14.5%]	11,070 [14.5%]	492 [15.3%]	4920 [15.3%]	266 [15.2%]	2660 [15.2%]

Data presented as mean [95% confidence interval] for continuous variables and number (%) for categorical variables.
AD Alzheimer's disease, *VD* vascular dementia.

time window before dementia diagnosis. The complete list of significant assessments and their detailed statistics were reported in the appendix for dementia (Supplementary eTables 3–16). We picked up the most representative assessments under each category and plotted their OR trajectories along the longitudinal timeframes preceding the diagnosis (Fig. 3). In general, as the time gets closer to the index date, dementia subjects witnessed worse functional performance or higher prevalence of health conditions since the fitted lines of OR gradually moved away from the reference level of 1.

Syndrome-related changes can be found over a decade preceding the diagnosis of dementia. According to the association results (Supplementary eTables 3–17, Fig. 3), cognitive functions witnessed significant declines in logical thinking (results score of the fluid intelligent test in 10–11 year OR: 0.79, 95% confidence interval [0.69–0.90], $p < 0.01$), pairs matching test (number of correct matches in 0–15 year OR: 0.84 [0.83–0.86], $p < 0.01$), prospective memory (results score in 11–12 year OR: 0.80 [0.71–0.89], $p < 0.01$), and reaction speed (time of correctly identifying matches in 11–12 year OR: 1.14 [1.08–1.21], $p < 0.01$) (Fig. 3A, Supplementary eTables 4 and 11). Physical measurements included declined strength in hand grip (0–15 year OR: 0.65 [0.63–0.67], $p < 0.01$), lung dysfunction (forced expiratory volume in 1-second, FEV1, in 0–15 year OR: 0.78 [0.76–0.81], $p < 0.01$), hearing impairment (speech-reception-threshold, SRT, in 10–11 year OR 1.18 [1.04–1.33], $p = 0.03$) (Fig. 3C, Supplementary eTables 8 and 15). Serum biomarkers of renal disorders (cystatin C in 0–15 year OR: 1.13 [1.11–1.16], $p < 0.01$), anemia (hemoglobin concentration in 10–11 year OR: 0.88 ([0.81–0.94], $p < 0.01$), and vitamin D deficiency (10–11 year OR: 0.88 ([0.81–0.94], $p < 0.01$) also emerged over 10 years before the dementia onset (Fig. 3E, Supplementary eTables 7 and 14). Complex mental disorders of long-term feelings and short-term mood fluctuations (Fig. 3D, Supplementary eTables 5 and 12), comorbidities of coronary heart disease (CHD), stroke, diabetes, depression, vision impairment (Fig. 3G, Supplementary eTables 9 and 16), as well as self-reported conditions of chest pain, falls in the last year and long-standing illness (Fig. 3H, Supplementary eTables 9 and 16) all exhibited higher prevalence over 10 years preceding the diagnosis of dementia.

Several associations exhibited short-term impact within six years preceding the dementia diagnosis. Such as inadequate activity (overall time of activity 3–5 year OR of 0.82 ([0.73–0.92], $p = 0.01$), drowsiness (sleep durations 3–5 year OR 1.13 [1.04–1.24], $p = 0.03$) (Fig. 3B, Supplementary eTables 3 and 10) and weight loss (weight in 0–5 year OR: 0.9 [0.83–0.98], $p = 0.02$) (Fig. 3C). Serum biomarkers of enriched endocrine (sex hormone-binding globulin, SHBG, 3–5 year OR: 1.20 ([1.09–1.31], $p < 0.01$), liver dysfunction (aspartate-amino transferase, AST, 3–5 year OR: 1.13 ([1.05–1.21], $p < 0.01$), infection risk (neutrophil percentage, 3–5 year OR: 1.30 ([1.18–1.43], $p < 0.01$) were found significant associations within 3–5 years (Fig. 3E, Supplementary eTables 7 and 14). Lipoprotein of small HDL concentration (6–7 years: OR 0.68 [0.55–0.85], $p < 0.01$) and large HDL concentration (0–5 years: OR 1.22 [1.06–1.41], $p = 0.02$) witnessed an opposite trend. In addition, enriched blood ketones of acetone (3–5 years: OR 1.24 [1.08–1.43], $p = 0.02$) and deficiency of brand-chain amino acids (BCAA) (0–5 years: OR 0.71 ([0.60–0.84], $p < 0.01$) and polyunsaturated fatty acids (PUFA) (0–5 years: OR 0.83 ([0.71–0.97], $p = 0.04$) were found associations within five years before the index date of dementia (Fig. 3F, Supplementary eTables 6 and 13). Comorbidities of hypertension, hearing loss, COPD, and arthritis demonstrated significantly higher prevalence in demented individuals in the similar prodromal time window (Fig. 3G, Supplementary eTables 9 and 16).

Pre-diagnostic functional changes in AD and VD

Compared to dementia, fewer assessments and health conditions were identified, 50 and 53 have surpassed the FDR correction, for AD and VD, respectively. Most selected predictors were pre-identified in the analyses of dementia; however, their year of significance preceding the index date varied (Supplementary eTables 18–31 for AD, Supplementary eTables 33–44 for VD, Fig. 4B, C). In general, fewer pre-diagnostic syndromes witnessed long-term associations over 10 years preceding the diagnosis of AD.

Specifically, predictors selected by both AD and VD included cognitive deficits in fluid intelligence (results score: 5–6 year in AD, 3–5 year in VD), pairs matching (number of correct matches: 8–9 year in AD, 3–5 year in VD), and reaction time (time to identify matches correctly: 9–10 year in AD; 12–15 year in VD) (Supplementary eTables 19, 26, 34 and 40); physical measurements of lung dysfunction (peak expiratory flow, PEF: 9–10 in year AD, 11–12 in year VD) and declined strength (11–12 year in AD, 11–12 year in VD) (Supplementary eTables 23, 30, 37 and 43); health conditions of CHD (3–5 year in AD; 12–15 year in VD) and self-reported falls in the last year (7–8 in AD, 6–7 in VD) (Supplementary eTables 24, 31, 38 and 44). In addition, mental disorders of nervous feelings and short-term mood fluctuations of depressed mood, tenseness, and unenthusiasm were also found with higher prevalence for both AD and VD individuals within five years before the diagnosis (Supplementary eTables 20, 27, 35 and 41). Noted, comorbid conditions of stroke, diabetes, hypertension, and self-reported chest pain, which were significant in VD, indicated no strong associations with AD (Supplementary eTables 24, 31, 38 and 44). Moreover, serum biomarkers of Apolipoprotein B, IGF-1, LDL direct were additionally found to have significant associations that emerged within the five years before the VD diagnosis, which were not pre-identified in the analysis of dementia (Supplementary eTables 22, 29, 36 and 42).

DISCUSSION

It was demonstrated in this large agnostic study that a broad range of clinical phenotypes were markedly associated with subsequent dementia in the years thereafter and could be considered as potential pre-diagnostic features. Associations noted that over a decade preceding dementia diagnosis comprised cognitive deficits, declined physical strength, increased falls, lung dysfunction, and complex mental problems. Comorbidities of CHD, stroke, diabetes, depression, vision impairment as well as serum biomarkers of renal disorders, anemia, and vitamin D deficiency also witness higher prevalence at similar times. This is followed by changes in body fat and lipoprotein metabolism as early as six to ten years prior to the onset of dementia. Within three to six years before dementia is diagnosed, activity inadequacy, sleep disturbances, hypertension, hearing loss, infection, endocrine and liver disorders, and dysregulations of ketone bodies started to manifest. Significant associations of deficiency in fatty acids and amino acids, COPD and arthritis were witnessed in the last stage before the index (Fig. 4A). All of these associations remained significant until the time of dementia diagnosis.

Traditional studies usually focused on exploring neuropathological hallmarks for the early detection of disease [1, 2, 11]; however, health-related changes in the pre-symptomatic phase may be underrecognized and underreported. We investigated an abundance of phenotypes on subsequent dementia risk, filling the gap in this field. In line with former findings [4, 6], cognitive decline emerged over a decade before the diagnosis of dementia. Specifically, memory loss came first, followed by prolonged reaction time, and declined logical thinking. In routine clinical practice, it's time-consuming to perform thorough and detailed cognitive testing. Our findings call for general practitioners to prioritize screening for these specific domains over general cognitive testing, where abnormalities tend to foreshadow possible disease progression to a comparatively advanced stage [4].

Apart from previously reported grip strength, increased falls, and weight change [6], we present evidences for pre-diagnostic changes in body fat, hearing tests, and lung function, which have

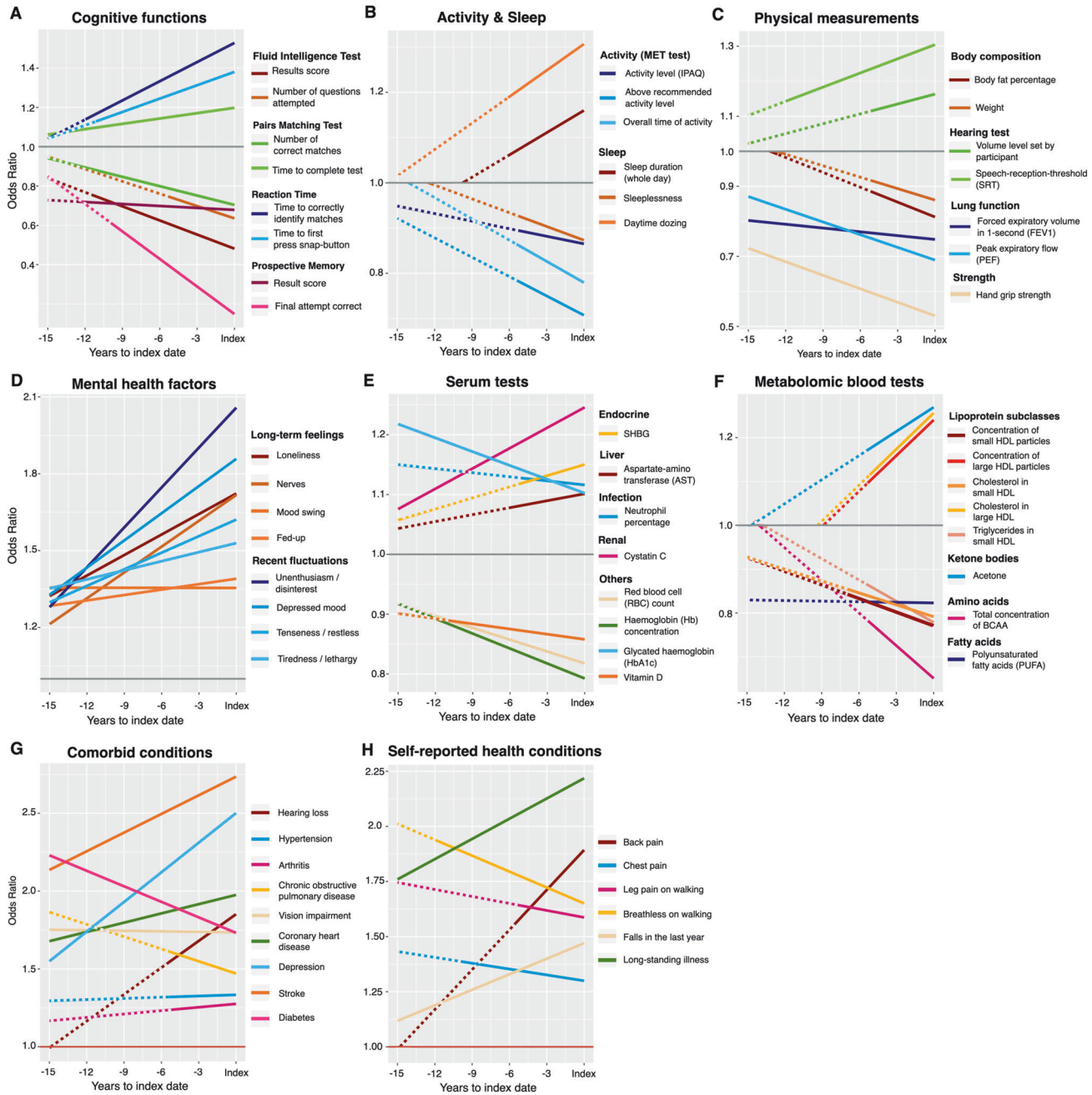

Fig. 3 Odds ratio trajectories of clinical assessments and health conditions of dementia. The trajectory lines of multi-domain clinical assessments and health conditions from (**A**) cognitive functions, (**B**) activity and sleep, (**C**) physical measurements, (**D**) mental health factors, (**E**) serum tests, (**F**) metabolomic blood tests, (**G**) comorbid conditions and (**H**) self-reported health conditions. The trajectory lines were fitted with linear regression models on odds ratios across time preceding the diagnosis of dementia. Odds ratios under each timeframe were derived from multivariate logistic regressions adjusted with age, gender, ethnicity, educational years, Townsend deprivation index, and durations from baseline to dementia index. Clinical assessments of continuous variables were normalized. Dotted lines indicate the assessments or health conditions were insignificant under specific timeframes. BCAA brand-chain amino acids, HDL high-density lipoprotein.

been continuously revealed [1, 9, 10, 12–15]. Metabolism, inflammation, and brain structures may contribute to explaining the relationship between body fat and dementia [12]. Hearing impairment is linked to greater temporal lobe atrophy [16]. No causal relationship was demonstrated between declined lung function and cognitive deficits, and the observed associations may be partially ascribed to their shared etiological factors [15, 17–19].

The development of dementia may also be related to physical inactivity and sleep disturbances, in agreement with the current consensus [1, 9, 10]. The pre-diagnostic linear risk changes in physical and behavioral measures support that they represent an early progressive loss of function rather than a low functional baseline. People may have impaired physical function and disrupted behaviors due to prodromal dementia, and these changes can be either a consequence or a cause of dementia, or both [1]. Maintaining physical function, enhancing exercise, and improving sleep should be sustained and continue nearer the time of risk.

Hearing loss was among the greatest risk factors for dementia [1]. It is worth noting that hearing tests became abnormal 11–12

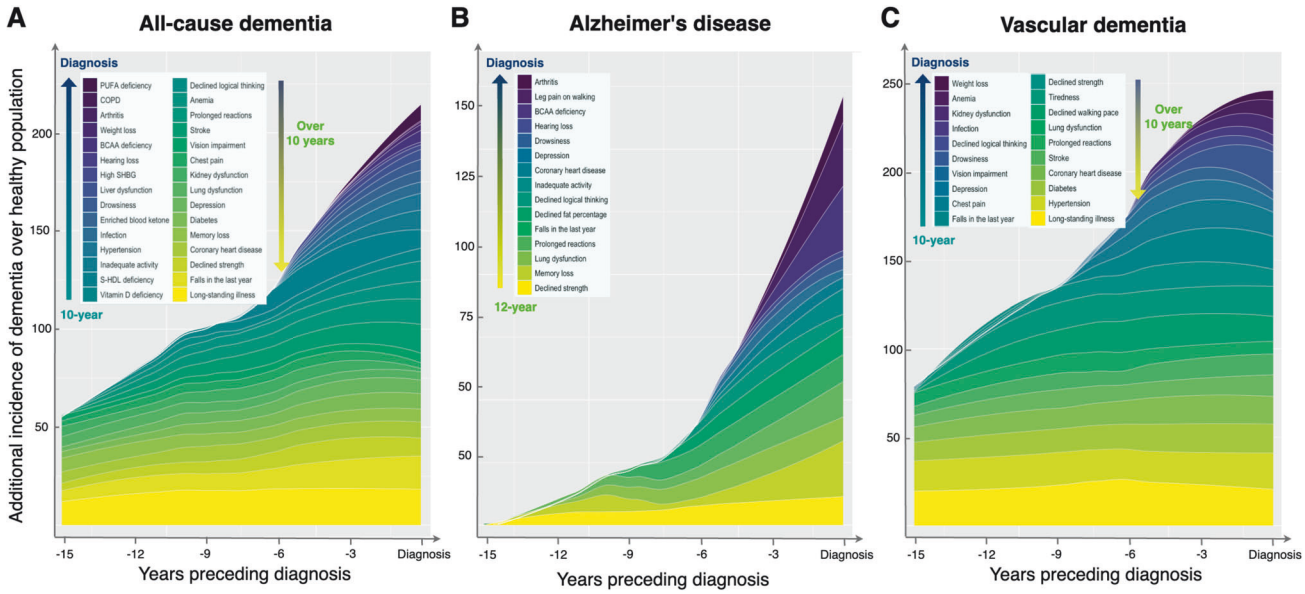

Fig. 4 Trajectories of health conditions preceding diagnosis of additional incidence rates of dementia, Alzheimer's disease, and vascular dementia over healthy controls. Stacked additional incidence curves of (**A**) all-cause dementia, (**B**) Alzheimer's disease, and (**C**) vascular dementia. Health conditions were selected from representative assessments or comorbid or self-reported conditions that were significantly associated with each outcome. The additional incidence was derived from the incidence of the target population minus that of the controls. The stacked incidence curves were plotted with statistics under ten timeframes preceding the dementia index, and any incidences under insignificant timeframes were set as zeros. BCAA brand-chain amino acids, COPD chronic obstructive pulmonary disease, PUFA polyunsaturated fatty acids, SHBG sex hormone-binding globulin, S-HDL-C small HDL concentration.

years before dementia diagnosis, whereas hearing loss was seen 3–5 years prior to the index, with obviously elevated risks compared to the former. Similar patterns were also found for lung function and COPD. The long-time interval provides critical early prevention opportunities for using hearing aids, which is proven to be protective for cognitive preservation [1]. Nonetheless, whether improvements in lung function mitigate the risk of dementia remains open [1, 9, 10]. Moreover, growing studies have linked visual impairment to dementia risk and treated it as a viable target for intervention [20, 21]. We reinforced this idea by finding that it could consistently increase dementia risk from 12 to 15 years before diagnosis.

Multimorbidity occurring before dementia was observed recently [8], and we went on to refine the prior findings by examining specific disease conditions. CHD, stroke, hypertension, and diabetes have been repeatedly proposed as potentially modifiable risk factors for dementia [1, 9, 10]. Consistent with this, our analysis further delineated their pre-diagnostic trajectories. We unraveled a novel relationship between arthritis and subsequent dementia, whereas a recent study suggested no association [8]. Different age ranges, study designs, and sample sizes may help explain the divergence. As for depression, whether it is a risk factor for dementia [1], a precursor or prodromal symptom of dementia [7], or both [3], remains a matter of debate. Future in-depth exploration of its role in dementia is necessary.

To our best knowledge, the pre-diagnostic trajectories of biochemical and metabolic factors have not been well-studied so far. Despite this, our results gain support from former studies showing associations between dementia and these broad blood indexes [22–25]. Biochemical markers reflecting systemic disorders became abnormal years before diagnosis and may be early warning signs of dementia. They could participate in the pathogenesis of dementia by affecting oxidative stress, inflammatory response, autophagy, and clearance of pathological hallmarks [22]. Regarding vitamin D supplementation, its benefits for dementia remain controversial [26]. This study, for the first time, disclosed several novel relationships before dementia diagnosis, such as the protective effect of small HDL-C but the inverse effect of large HDL-C. Individuals with higher BCAA or PUFA were protected against dementia, in line with former findings [24, 27]. Randomized clinical trials, however, negated the effect of PUFA on neurocognitive outcomes [28]. Our Findings on ketone bodies disagreed with the cognitive protective effects of ketogenic diets [29, 30]. Whether the mechanisms involved in the development of dementia differ between a passive high ketone state and active ketogenesis requires future exploration.

The observed higher prevalence of complex mental disorders characterized by long-term feelings and short-term mood fluctuations preceding dementia diagnosis prompts a critical examination of mental health factors within the context of Minimal Behavioral Impairment (MBI). In future observational cohorts, employing a structured approach like the MBI checklist for assessing mood and behavioral symptoms would be prudent. This systematic tool enables the comprehensive evaluation of subtle behavioral changes associated with dementia risk, facilitating early detection and intervention. By standardizing assessment processes and quantifying symptoms, researchers can better understand the progression of cognitive decline and inform targeted interventions to improve patient outcomes.

When comparing the pre-diagnostic functional changes between AD and VD, our findings reveal both expected commonalities and notable distinctions. As AD constitutes the majority of dementia [31], it is unsurprising that pre-diagnostic functional changes in AD and dementia largely overlap. Cognitive deficits in fluid intelligence, pairs matching, and reaction time were identified as shared predictors, albeit with variations in the timing of their significance preceding the index date. However, our comparison with VD revealed distinct patterns. Pre-diagnostic functional changes in VD predominantly manifested as vascular comorbidities, aligning with the close association between vascular pathology and VD [32]. This finding is substantiated by the distinct etiological pathways underlying AD and VD [33]. While shared predictors such as cognitive deficits and physical measurements were identified for both AD and VD, their temporal

significance varied, suggesting nuanced disease progression trajectories. Notably, mental disorders exhibited higher prevalence in the pre-diagnostic phase for both AD and VD, suggesting potential shared mechanisms in disease manifestation. Conversely, certain comorbid conditions and serum biomarkers, significant in VD, showed no strong associations with AD, underscoring differential disease pathophysiology. These findings underscore the complex interplay between different factors in the pre-diagnostic phase of AD and VD, offering insights into potential avenues for early detection and intervention tailored to the specific pathology of each condition [32].

Collecting data on all antecedent characteristics is not feasible in real-world clinical practice. Our findings lend vital evidence to the priority screening that primary care physicians ought to undertake and the opportunity to resolve patient concerns in the early pre-diagnostic phase. Rather than identifying asymptomatic patients, this is about accurately pinpointing the potential cause of patients exhibiting symptoms or signs and then seeking prompt referrals. Otherwise, patients may have to wait many years or even decades for their symptoms to be explained. By then, the optimal time to slow or halt disease progression has perhaps been missed. Current strategies targeted at improving cognition might have been initiated too late relative to disease progression, and a wide scope of early interventions and treatments is critical and necessary.

The strength of this study lies in the comprehensive and in-depth investigations of health-related phenotypes through a large-scale, exploratory, data-driven approach. Besides, the prolonged follow-up allowed us to track trajectories for up to 15 years prior to dementia diagnosis. Some caveats should also be acknowledged. First, the UK Biobank is skewed toward those at lower risk of disease and lacks representativeness of racial and socioeconomic diversity; thus, conclusions drawn in our analysis still require further validation. Second, this study used longitudinal measure of dementia diagnosis while the clinical assessments were collected at baseline; therefore, we performed a nested case-control strategy to reform the data structure for longitudinal interpretation of clinical trajectories. However, further studies utilizing longitudinal measures of clinical variables are necessary to confirm and validate our findings. Third, the age range of the population recruited to the UK Biobank was 40 to 90 years, potentially ruling out the age at which some factors play the strongest role. Fourth, the absence of detailed diagnostic methods, such as lumbar punctures or imaging techniques employed by doctors, prevents the direct alignment of UK Biobank data with the current ATN biological definition of AD and updates on VD. This impedes a comprehensive understanding of the biological underpinnings of these conditions within the context of the database.

This is the first study to examine such a wide range of potential risk factors and early manifestations of dementia and its subtypes of AD and VD. The study determines when these factors appear in the disease process up to 15 years before the first diagnosis. Our findings present a comprehensive temporal-diagnostic landscape preceding incident dementia, offering novel insights into early prevention and disease-modifying therapies.

DATA AVAILABILITY

All data used in this study were accessed from the publicly available UK Biobank Resource under application number 19542. These data cannot be shared with other investigators.

CODE AVAILABILITY

All analyses were performed with the StatsModel (v0.11.1) and ScikitLearn (v0.24.1) packages in Python (v3.6). The code can be accessed through https://github.com/jasonHKU0907/DementiaPhenoTrajectory/.

REFERENCES

1. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet. 2020;396:413–46.
2. Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, et al. Defeating Alzheimer's disease and other dementias: a priority for European science and society. Lancet Neurol. 2016;15:455–532.
3. Nedelec T, Couvy-Duchesne B, Monnet F, Daly T, Ansart M, Gantzer L, et al. Identifying health conditions associated with Alzheimer's disease up to 15 years before diagnosis: an agnostic study of French and British health records. Lancet Digit Health. 2022;4:e169–e178.
4. Bateman RJ, Xiong C, Benzinger TL, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med. 2012;367:795–804.
5. Schermer MHN, Richard E. On the reconceptualization of Alzheimer's disease. Bioethics. 2019;33:138–45.
6. Swaddiwudhipong N, Whiteside DJ, Hezemans FH, Street D, Rowe JB, Rittman T. Pre-diagnostic cognitive and functional impairment in multiple sporadic neurodegenerative diseases. Alzheimers Dement. 2022;19:1752–63.
7. Singh-Manoux A, Dugravot A, Fournier A, Abell J, Ebmeier K, Kivimaki M, et al. Trajectories of Depressive Symptoms Before Diagnosis of Dementia: A 28-Year Follow-up Study. JAMA Psychiatry. 2017;74:712–8.
8. Chen H, Zhou Y, Huang L, Xu X, Yuan C. Multimorbidity burden and developmental trajectory in relation to later-life dementia: A prospective study. Alzheimers Dement. 2023;19:2024–33.
9. Yu JT, Xu W, Tan CC, Andrieu S, Suckling J, Evangelou E, et al. Evidence-based prevention of Alzheimer's disease: systematic review and meta-analysis of 243 observational prospective studies and 153 randomised controlled trials. J Neurol Neurosurg Psychiatry. 2020;91:1201–9.
10. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet. 2017;390:2673–734.
11. Jack CR Jr. Advances in Alzheimer's disease research over the past two decades. Lancet Neurol. 2022;21:866–9.
12. Deng YT, Li YZ, Huang SY, Ou YN, Zhang W, Chen SD, et al. Association of life course adiposity with risk of incident dementia: a prospective cohort study of 322,336 participants. Mol Psychiatry. 2022;27:3385–95.
13. Kuo K, Zhang YR, Chen SD, He XY, Huang SY, Wu BS, et al. Associations of grip strength, walking pace, and the risk of incident dementia: A prospective cohort study of 340212 participants. Alzheimers Dement. 2023;19:1415–27.
14. Loughrey DG, Kelly ME, Kelley GA, Brennan S, Lawlor BA. Association of Age-Related Hearing Loss With Cognitive Function, Cognitive Impairment, and Dementia: A Systematic Review and Meta-analysis. JAMA Otolaryngol Head Neck Surg. 2018;144:115–26.
15. Lutsey PL, Chen N, Mirabelli MC, Lakshminarayan K, Knopman DS, Vossel KA, et al. Impaired Lung Function, Lung Disease, and Risk of Incident Dementia. Am J Respir Crit Care Med. 2019;199:1385–96.
16. Armstrong NM, An Y, Doshi J, Erus G, Ferrucci L, Davatzikos C, et al. Association of Midlife Hearing Impairment With Late-Life Temporal Lobe Volume Loss. JAMA Otolaryngol Head Neck Surg. 2019;145:794–802.
17. Russ TC, Harris SE, Batty GD. Pulmonary Function and Risk of Alzheimer Dementia: Two-Sample Mendelian Randomization Study. Chest. 2021;160:274–6.
18. Higbee DH, Granell R, Hemani G, Smith GD, Dodd JW. Lung function, COPD and cognitive function: a multivariable and two sample Mendelian randomization study. BMC Pulm Med. 2021;21:246.
19. Higbee D, Granell R, Walton E, Korologou-Linden R, Davey Smith G, Dodd J. Examining the possible causal relationship between lung function, COPD and Alzheimer's disease: a Mendelian randomisation study. BMJ Open Respir Res. 2021;8:e000759.
20. Ehrlich JR, Goldstein J, Swenor BK, Whitson H, Langa KM, Veliz P. Addition of Vision Impairment to a Life-Course Model of Potentially Modifiable Dementia Risk Factors in the US. JAMA Neurol. 2022;79:623–6.
21. Shang X, Zhu Z, Wang W, Ha J, He M. The Association between Vision Impairment and Incidence of Dementia and Cognitive Impairment: A Systematic Review and Meta-analysis. Ophthalmology. 2021;128:1135–49.
22. He XY, Kuo K, Yang L, Zhang YR, Wu BS, Chen SD, et al. Serum clinical laboratory tests and risk of incident dementia: a prospective cohort study of 407,190 individuals. Transl Psychiatry. 2022;12:312.
23. van der Lee SJ, Teunissen CE, Pool R, Shipley MJ, Teumer A, Chouraki V, et al. Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies. Alzheimers Dement. 2018;14:707–22.
24. Tynkkynen J, Chouraki V, van der Lee SJ, Hernesniemi J, Yang Q, Li S, et al. Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: A prospective study in eight cohorts. Alzheimers Dement. 2018;14:723–33.

25. Chouraki V, Preis SR, Yang Q, Beiser A, Li S, Larson MG, et al. Association of amine biomarkers with incident dementia and Alzheimer's disease in the Framingham Study. Alzheimers Dement. 2017;13:1327–36.
26. Liu D, Meng X, Tian Q, Cao W, Fan X, Wu L, et al. Vitamin D and Multiple Health Outcomes: An Umbrella Review of Observational Studies, Randomized Controlled Trials, and Mendelian Randomization Studies. Adv Nutr. 2022;13:1044–62.
27. Devassy JG, Leng S, Gabbs M, Monirujjaman M, Aukema HM. Omega-3 Polyunsaturated Fatty Acids and Oxylipins in Neuroinflammation and Management of Alzheimer Disease. Adv Nutr. 2016;7:905–16.
28. Brainard JS, Jimoh OF, Deane KHO, Biswas P, Donaldson D, Maas K, et al. Omega-3, Omega-6, and Polyunsaturated Fat for Cognition: Systematic Review and Meta-analysis of Randomized Trials. J Am Med Dir Assoc. 2020;21:1439–50.e1421.
29. Rusek M, Pluta R, Ulamek-Koziol M, Czuczwar SJ. Ketogenic Diet in Alzheimer's Disease. Int J Mol Sci. 2019;20:3892.
30. Jennings A, Cunnane SC, Minihane AM. Can nutrition support healthy cognitive ageing and reduce dementia risk? BMJ. 2020;369:m2269.
31. 2023 Alzheimer's disease facts and figures. Alzheimers Dement. 2023;19:1598–695.
32. Kalaria RN. The pathology and pathophysiology of vascular dementia. Neuropharmacology. 2018;134:226–39.
33. O'Brien JT, Thomas A. Vascular dementia. Lancet. 2015;386:1698–706.

ACKNOWLEDGEMENTS

This study utilized the UK Biobank Resource under application number 19542. We want to thank all the participants and researchers from the UK Biobank. This study was funded by grants from the National Key Research and Development Program of China (grant no. 2023YFC3605400), Science and Technology Innovation 2030 Major Projects (2022ZD0211600), Shanghai Pujiang Program (23PJD006), National Key R&D Program of China (2018YFC1312904, 2019YFA0709502), National Natural Science Foundation of China (82071201, 81971032, 82071997). Further, we would like to thank the support from the Shanghai Municipal Science and Technology Major Project (2018SHZDZX01), Research Start-up Fund of Huashan Hospital (2022QD002), Excellence 2025 Talent Cultivation Program at Fudan University (3030277001), Shanghai Rising-Star Program (21QA1408700), the 111 Project (No. B18015) and Shanghai Center for Brain Science and Brain-Inspired Technology, ZHANGJIANG LAB, Tianqiao and Chrissy Chen Institute, and the State Key Laboratory of Neurobiology and Frontiers Center for Brain Science of Ministry of Education, Fudan University.

AUTHOR CONTRIBUTIONS

WC, JTY, and JFF conceived, designed, and supervised the project. JY implemented statistical analyses. YG, YJW, YZ, HFW, LBW, and JJK supported the study and contributed to the discussion of the results. JY and YG drafted the manuscript and accessed and verified the underlying data reported in the manuscript. investigator was blinded to the group allocation during the experiment and/or when assessing the outcome. All authors had full access to all the study data and accepted the responsibility to submit it for publication.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41380-024-02570-0.

Correspondence and requests for materials should be addressed to Jian-Feng Feng, Jin-Tai Yu or Wei Cheng.