Specific midlife depressive symptoms and long-term dementia risk: a 23-year UK prospective cohort study

Philipp Frank, Archana Singh-Manoux, Jaana Pentti, G David Batty, Andrew Sommerlad, Andrew Steptoe, Gill Livingston, Robert Howard, Mika Kivimäki

Lancet Psychiatry 2026; 13: 100–11

Published Online December 15, 2025 https://doi.org/10.1016/S2215-0366(25)00331-1

See Comment page 88

Division of Psychiatry, Brain Sciences, University College London, London, UK (P Frank PhD, Prof A Singh-Manoux PhD, Prof A Sommerlad PhD, Prof G Livingston MD PhD, Prof R Howard MD, Prof M Kivimäki PhD FMedSci); **Clinicum, University of Helsinki, Helsinki, Finland** (J Pentti BSc, Prof M Kivimäki); **Research Department of Epidemiology and Public Health, University College London, London, UK** (Prof G D Batty DSc FRSE); **Université Paris Cité, Inserm U1153, Paris, France** (Prof A Singh-Manoux); **North London NHS Foundation Trust, London, UK** (Prof A Sommerlad, Prof G Livingston, Prof R Howard); **Department of Behavioural Science and Health, University College London, London, UK** (Prof A Steptoe PhD FMedSci)

Correspondence to: Dr Philipp Frank, Division of Psychiatry, Brain Sciences, University College London, London W1T 7NF, UK **philipp.frank.16@ucl.ac.uk**

Summary

Background Midlife depression has been associated with an increased risk of dementia, but it remains unclear whether this risk is attributable to specific symptoms. We aimed to identify the midlife depressive symptoms most strongly linked to subsequent dementia and to ascertain whether these associations were independent of established dementia risk factors.

Methods In this prospective, observational cohort study based on the UK Whitehall II study, participants (aged 35–55 years at study inception [1985–88]) were eligible for analysis if they had complete depression data and successful linkage to national health records; individuals with prevalent dementia at baseline were excluded. In 1997–99, the baseline for this analysis, participants underwent a clinical examination and completed the 30-item version of the General Health Questionnaire (GHQ-30, a validated screening instrument for detecting clinically significant psychiatric distress in the general population). Threshold-level depression was defined as a GHQ-30 score of 5 or higher. The primary outcome was incident dementia, ascertained via linkage to UK National Health Service (NHS) Hospital Episode Statistics for inpatient admissions, the Mental Health Services Data Set, or the NHS Central Registry for mortality from April 24, 1997, to March 1, 2023. Analyses were conducted using a series of multivariable-adjusted Cox proportional hazards regression models. Hazard ratios (HRs) and accompanying 95% CIs were adjusted for age, sex, and ethnicity in the basic model. People with lived experience were not involved in the study design or writing process.

Findings Of 6511 participants in the Whitehall II study who completed the GHQ-30 between April 24, 1997, and Jan 8, 1999, 5811 were eligible for this analysis. The mean age of participants was 55·7 years (SD 6·0; range 45–69); 1646 (28·3%) participants were women, 4165 (71·7%) were men, 5356 (92·2%) reported their ethnicity as White, and 455 (7·8%) reported their ethnicity as non-White. During a mean follow-up of 22·6 years (SD 5·0), 586 participants (10·1%) developed dementia. Six depressive symptoms emerged as robust midlife indicators of increased dementia risk: "Losing confidence in myself" (HR 1·51, 95% CI 1·16–1·96), "Not able to face up to problems" (1·49, 1·09–2·04), "Not feeling warmth and affection for others" (1·44, 1·06–1·95), "Nervous and strung-up all the time" (1·34, 1·03–1·72), "Not satisfied with the way tasks are carried out" (1·33, 1·05–1·69), and "Difficulties concentrating" (1·29, 1·01–1·65). Associations were independent of established dementia risk factors, including *APOEε4* status, cardiometabolic conditions, and lifestyle factors. In individuals younger than 60 years at baseline, the six symptoms fully accounted for the association between midlife depression and dementia risk.

Interpretation A distinct set of midlife depressive symptoms was associated with an increased risk of dementia, suggesting that these symptoms might be early markers of underlying neurodegenerative processes. These findings could inform earlier identification and more targeted interventions for individuals with depression who are at risk of dementia.

Funding Wellcome Trust, UK Medical Research Council, National Institute on Aging, and Research Council of Finland.

Introduction

With rapidly ageing populations, dementia has emerged as a major global public health concern. Dementia, with Alzheimer's disease accounting for approximately 70% of cases, currently affects more than 55 million people worldwide, a figure expected to reach 153 million by 2050.[1] Identifying and managing modifiable risk factors in midlife (ie, age 40–60 years) to delay the onset of or prevent dementia has, therefore, become a global public health priority.[2]

The 2024 *Lancet* Commission report on dementia prevention, intervention, and care highlights depression—the second leading cause of disability globally[3]—as a potentially modifiable midlife risk factor.[4] Although depression is highly comorbid with dementia,[5] and there is consistent evidence linking later-life

Research in context

Evidence before this study

Growing evidence indicates that midlife depression might increase the risk of subsequent dementia. We searched PubMed from database inception to July 10, 2025, using the terms "depression", "depressive symptoms", and "dementia", without language or date restrictions, and we considered original research articles, systematic reviews, and meta-analyses. Most studies conceptualised depression as a binary exposure and focused on later-life depression. Although depression is a clinically heterogeneous condition with varying types of symptom expressions, no large-scale longitudinal study has systematically examined whether specific depressive symptoms underlie the association between midlife depression and later-life dementia.

Added value of this study

This prospective cohort study identified six midlife depressive symptoms that were robustly associated with increased dementia risk over a 23-year follow-up, including loss of confidence, concentration difficulties, reduced problem-solving abilities, impaired social connections, and persistent nervousness. These associations were largely independent of established dementia risk factors and fully accounted for the association between depression in midlife (<60 years) and subsequent risk of dementia. Moving beyond a binary conceptualisation of depression, our findings extend previous work by highlighting symptom-specific pathways and showing that a subset of depressive features appear to underlie the long-term risk of dementia.

Implications of all the available evidence

Our findings suggest that not all depressive symptoms contribute equally to dementia risk. Focusing on specific symptom patterns rather than treating depression as a single, unitary construct could improve early identification of individuals at increased risk of dementia and inform targeted prevention strategies. Further research is needed to test these findings in more diverse populations and explore the biological and behavioural mechanisms linking specific symptom profiles to dementia. In clinical practice, if these findings are validated, routine assessment of symptom profiles could support earlier dementia risk stratification and more personalised, multimodal treatment approaches.

depression with an increased dementia risk,[6] the role of midlife depression in dementia pathology remains poorly understood.[6–8] Few studies have sufficiently long follow-up periods to track participants from midlife into later life, and among those that have such follow-up periods, findings have been mixed.[6,7,9]

To advance our understanding of how midlife depression might relate to dementia risk, it is important to move beyond viewing depression as a single entity. Depression is a clinically heterogeneous condition—a syndrome encompassing more than 1000 possible symptom combinations that meet the diagnostic criteria of major depressive disorder.[10,11] It has been hypothesised that treating the diverse presentations of depression as a unitary construct could have contributed to inconsistencies in the literature,[11–12] hindering the identification of subgroups most at risk of developing dementia. A more nuanced understanding of symptom-specific associations might help clinicians to distinguish between middle-aged patients whose depression reflects an elevated dementia risk and those whose symptoms are more likely to be due to other causes, thereby informing clinical evaluation and tailoring of treatment. To date, no studies have systematically examined the association between midlife depression and dementia risk at the level of individual depression symptoms.

To address this knowledge gap, we used longitudinal data from a large, prospective cohort of adults to examine the link between individual midlife depressive symptoms and subsequent dementia risk over a mean follow-up period of 23 years.

Methods

Study design and participants

This study is a secondary analysis of the UK Whitehall II study,[13] a prospective cohort study of men and women aged 35–55 years recruited from 20 London-based civil service departments between 1985 and 1988. The Whitehall II study—also referred to as the Stress and Health study—was initially established to explore determinants of social inequalities in health. Individuals were eligible for the present study if they participated in a medical examination, including an assessment of depressive symptoms among other clinical and questionnaire-based measures, between 1997 and 1999 (appendix p 1). Participants with missing data on depression, clinical measures, or linkage data were excluded. Written informed consent was obtained from participants at each contact, and ethical approval for the study was granted by the University College London (UCL) Medical School Committee on the Ethics of Human Research (reference number 85/0938).

See Online for appendix

Measures

Assessments of depression were conducted between 1997 and 1999, using the 30-item version of the General Health Questionnaire (GHQ-30),[14] a self-reported validated screening instrument to detect clinically significant psychiatric distress, primarily symptoms of depression, in general population settings. The questionnaire was mailed to participants' homes before their clinical examination at the UCL Whitehall II clinic, to be completed in advance; responses were then

reviewed during the clinical examination. Participants were asked to indicate whether and how often each symptom had occurred during the past 2 weeks. Response options for each item were: not at all, no more than usual, rather more than usual, and much more than usual. The standard binary scoring method was applied, with the first two response options scored as 0 (ie, absence of symptom) and the latter two options as 1 (ie, presence of symptom), resulting in a total score ranging from 0 to 30. Threshold-level depression was defined as a total score of 5 or more.[15] In this cohort, the GHQ-30 has shown a good diagnostic performance when benchmarked against the interviewer-administered revised Clinical Interview Schedule, with sensitivity and specificity estimates of 78% and 83%, respectively, for detecting depressive episodes.[16]

Our analyses included a series of covariates captured at baseline (1997–99). Demographic characteristics included sex (male *vs* female), age, and ethnicity (White *vs* non-White). Ethnic origin was self-reported, with participants selecting from the following categories: Asian or Asian British, Black or Black British, Chinese, White, and Other ethnic group. Because most participants identified as White, there were insufficient dementia events within individual minority groups to support meaningful subgroup analysis; therefore, we categorised ethnicity as either White or non-White. *APOEε4* status, established using a standard PCR assay with DNA extracted from blood samples, was used as an indicator of genetic risk (0 *vs* 1 *vs* 2 alleles). Baseline age, sex, ethnicity, and *APOEε4* status were considered as potential confounders as these characteristics were ascertained before the onset of depressive symptoms and dementia.

We also assessed the role of ten *Lancet* Commission risk factors that were available in the Whitehall II study at baseline (1997–99):[4] smoking (current smoker *vs* never or ex-smoker), excessive alcohol use (≥21 units per week[17]), diabetes (one of the following: fasting glucose ≥7·0 mmol/L, non-fasting glucose 2h ≥11·1 mmol/L, antidiabetic drug use, or self-reported doctor diagnosis), elevated LDL cholesterol (≥130 mg/dL), hypertension (one of the following: systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, antihypertensive drug use, or self-reported doctor diagnosis), obesity (BMI ≥30 kg/m²), lower educational qualification (elementary or secondary education only), physical inactivity (WHO criteria[18] for engaging in moderate-to-vigorous physical activity for at least 2·5 h per week not met), social isolation (never or almost never visited by friends or family and living alone), and hearing impairment (assessed in Whitehall II in terms of difficulties following a conversation if there was background noise or difficulties hearing someone talking in a quiet room[4]). Because the direction of association between depressive symptoms and these risk factors cannot be ascertained in observational data, they might act as confounders, mediators, or both.

We examined cognitive function at baseline (1997–99) and at up to four subsequent clinic phases (2002–04, 2007–09, 2012–13, and 2015–16) as part of the routine Whitehall II clinical follow-up. These assessments allowed us to test whether individuals with dementia-related depressive symptoms at baseline had poorer mean cognitive performance across repeated clinical visits. Cognitive function was measured using a standardised test battery covering four domains:[19] memory (20-word free recall test); reasoning (Alice Heim 4-I test); phonemic fluency (words starting with s); and semantic fluency (animal names). Tests were administered in person at the UCL Whitehall II clinic. For each cognitive domain, scores at follow-up visits were standardised to the baseline mean and SD to create Z scores (mean of 0 and SD of 1). These domain-specific scores were then used to derive a global cognitive score at each phase by summing the domain scores and restandardising the total to the baseline distribution. This approach minimises the measurement error inherent in individual tests.

The primary outcome was incident dementia, which was measured after baseline. Dementia diagnoses and corresponding dates were ascertained from three UK National Health Service (NHS) sources, from April 24, 1997, to March 1, 2023: Hospital Episode Statistics inpatient admissions, the Mental Health Services Data Set (MHSDS), and the NHS Central Registry for mortality. Dementia was defined with WHO ICD-10, codes F00, F01, F03, G30, and G31 in Hospital Episode Statistics and the NHS Central Registry for mortality or as a dementia diagnosis recorded in MHSDS diagnostic fields by specialist mental health services. The NHS provides near-complete health-care coverage for all individuals legally residing in the UK. Participants with dementia at or before baseline were excluded from the analysis. Although hospital diagnoses of dementia are rarely validated with biomarkers beyond structural neuroimaging, the sensitivity and specificity of all-cause dementia ascertainment based on NHS data are high, at 0·78 and 0·92, respectively.[20]

Statistical analysis

Differences in baseline characteristics by depression and dementia status were examined using means and proportions, with statistical comparisons done with χ^2 tests for categorical variables and *t* tests for continuous variables. We estimated pairwise associations between all GHQ-30 items using ϕ coefficients. Hierarchical clustering was applied to identify symptom groupings based on co-occurrence patterns.

After confirming that the proportional hazards assumption had not been violated (appendix pp 2–3), we conducted separate Cox proportional hazards regression models to examine the associations of each GHQ symptom, as well as threshold-level depression (GHQ-30 score ≥5), with incident dementia. Hazard ratios (HRs)

and accompanying 95% CIs were adjusted for age, sex, and ethnicity (basic model). Symptoms that were significantly associated with incident dementia at conventional levels of statistical significance ($p<0·05$) were retained for further analyses.

We examined the structure of inter-relationships among depressive symptoms that were associated with incident dementia by estimating a regularised partial correlation network using graphical least absolute shrinkage and selection operator. This approach computes partial correlation coefficients between each pair of symptoms while adjusting for all others in the network. To test the robustness of observed symptom–dementia associations, we conducted a series of

	Full sample (N=5811)	Participants without dementia (N=5225)	Participants with dementia (N=586)	p value	Participants without depression (N=4563)	Participants with depression (N=1248)	p value
Age, years	55·7 (6·0)	55·2 (5·9)	60·6 (5·0)	<0·0001	56·2 (6·0)	54·0 (5·6)	<0·0001
Sex							
Men	4165 (71·7%)	3776 (72·3%)	389 (66·4%)	0·0027	3345 (73·3%)	820 (65·7%)	<0·0001
Women	1646 (28·3%)	1449 (27·7%)	197 (33·6%)	..	1218 (26·7%)	428 (34·3%)	..
Ethnicity							
White	5356 (92·2%)	4837 (92·6%)	519 (88·6%)	0·0006	4214 (92·4%)	1142 (91·5%)	0·32
Non-White	455 (7·8%)	388 (7·4%)	67 (11·4%)	..	349 (7·6%)	106 (8·5%)	..
Less education							
Yes	2081/5736 (36·3%)	1807/5157 (35·0%)	274/579 (47·3%)	<0·0001	1650/4505 (36·6%)	431/1231 (35·0%)	0·30
No	3655/5736 (63·7%)	3350/5157 (65·0%)	305/579 (52·7%)	..	2855/4505 (63·4%)	800/1231 (65·0%)	..
Missing	75	68	7	..	58	17	..
Smoking							
Yes	567/5782 (9·8%)	512/5198 (9·8%)	55/584 (9·4%)	0·74	451/4542 (9·9%)	116/1240 (9·4%)	0·55
No	5215/5782 (90·2%)	4686/5198 (90·2%)	529/584 (90·6%)	..	4091/4542 (90·1%)	1124/1240 (90·6%)	..
Missing	29	27	2	..	21	8	..
Excessive alcohol							
Yes	1349/5746 (23·5%)	1225/5167 (23·7%)	124/579 (21·4%)	0·22	1047/4512 (23·2%)	302/1234 (24·5%)	0·35
No	4397/5746 (76·5%)	3942/5167 (76·3%)	455/579 (78·6%)	..	3465/4512 (76·8)	932/1234 (75·5%)	..
Missing	65	58	7	..	51	14	..
Diabetes							
Yes	167/4916 (3·4%)	132/4428 (3·0%)	35/488 (7·2%)	<0·0001	135/3885 (3·5%)	32/1031 (3·1%)	0·56
No	4749/4916 (96·6%)	4296/4428 (97·0%)	453/488 (92·8%)	..	3750/3885 (96·5%)	999/1031 (96·9%)	..
Missing	895	797	98	..	678	217	..
Obesity							
Yes	702/5053 (13·9%)	621/4535 (13·7%)	81/518 (15·6%)	0·23	521/3953 (13·2%)	181/1100 (16·5%)	0·0055
No	4351/5053 (86·1%)	3914/4535 (86·3%)	437/518 (84·4%)	..	3432/3953 (86·8%)	919/1100 (83·5%)	..
Missing	758	690	68	..	610	148	..
Social isolation							
Yes	49/5205 (0·9%)	44/4705 (0·9%)	5/500 (1·0%)	0·81	40/4080 (1·0%)	9/1125 (0·08%)	0·58
No	5156/5205 (99·1%)	4661/4705 (99·1%)	495/500 (99·0%)	..	4040/4080 (99·0%)	1116/1125 (99·2%)	..
Missing	606	520	86	..	483	123	..
Hypertension							
Yes	1660/5796 (28·6%)	1441/5211 (27·7%)	219/585 (37·4%)	<0·0001	1363/4552 (29·9%)	297/1244 (23·9%)	<0·0001
No	4136/5796 (71·4%)	3770/5211 (72·3%)	366/585 (62·6%)	..	3189/4552 (70·1%)	947/1244 (76·1%)	..
Missing	15	14	1	..	11	4	..
LDL cholesterol							
Elevated	3515/5088 (69·1%)	3131/4585 (68·3%)	384/503 (76·3%)	0·0002	2788/3998 (69·7%)	727/1090 (66·7%)	0·054
Normal	1573/5088 (30·9%)	1454/4585 (31·7%)	119/503 (23·7%)	..	1210/3998 (30·3%)	363/1090 (33·3%)	..
Missing	723	640	83	..	565	158	..
Hearing loss							
Yes	1183/5748 (20·6%)	1033/5173 (20·0%)	150/575 (26·1%)	0·0006	817/4516 (18·1%)	366/1232 (29·7%)	<0·0001
No	4565/5748 (79·4%)	4140/5173 (80·0%)	425/575 (73·9%)	..	3699/4516 (81·9%)	866/1232 (70·3%)	..
Missing	63	52	11	..	47	16	..

(Table continues on next page)

	Full sample (N=5811)	Participants without dementia (N=5225)	Participants with dementia (N=586)	p value	Participants without depression (N=4563)	Participants with depression (N=1248)	p value
(Continued from previous page)							
Physical inactivity							
Yes	1754/5775 (30·4%)	1558/5191 (30·0%)	196/584 (33·6%)	0·077	1303/4536 (28·7%)	451/1239 (36·4%)	<0·0001
No	4021/5775 (69·6%)	3633/5191 (70·0%)	388/584 (66·4%)	..	3233/4536 (71·3%)	788/1239 (63·6%)	..
Missing	36	34	2	..	27	9	..
APOEε4 alleles							
0	3238/4482 (72·2%)	3012/4059 (74·2%)	226/423 (53·4%)	<0·0001	2556/3532 (72·4%)	682/950 (71·8%)	0·93
1	1139/4482 (25·4%)	974/4059 (24·0%)	165/423 (39·0%)	..	893/3532 (25·3%)	246/950 (25·9%)	..
2	105/4482 (2·3%)	73/4059 (1·8%)	32/423 (7·6%)	..	83/3532 (2·3%)	22/950 (2·3%)	..
Missing	1329	1166	163	..	1031	298	..

Data are n, n (%), n/N (%), or mean (SD).

Table: **Baseline characteristics of the study population by baseline depression status and dementia status at follow-up**

additional Cox proportional hazards regression models, assessing whether other known dementia risk factors could account for these associations. Analyses were conducted in the full sample and stratified by baseline age group (<60 years *vs* ≥60 years) to explore potential differences in midlife and later-life depressive symptoms.[6] All effect estimates were adjusted for baseline age, sex, and ethnicity (basic model). In subsequent models, each of the 11 established dementia risk factors (including *APOEε4* status) was added separately and, in a final model, all risk factors were combined (adjusted model). We calculated the percentage of HR attenuation after each adjustment using the following formula: $[(HR_{basic\ model} - HR_{adjusted\ model})/(HR_{basic\ model} - 1)] \times 100$. Missing data for baseline risk factors were imputed using multiple imputation by chained equations (20 imputations), assuming data were missing at random conditional on observed data. Age, sex, and ethnicity were treated as fully observed, whereas follow-up time and dementia (outcome) were included as auxiliary variables. Estimates were pooled across imputations using Rubin's rules.[21]

To estimate the cumulative probability of participants remaining dementia free over the follow-up period, we generated Kaplan–Meier survival curves stratified by age at baseline (<60 years *vs* ≥60 years). Survival curves were plotted for individuals with at least one of the dementia-related depressive symptoms versus those without dementia-related depressive symptoms; survival curves were also plotted separately for each symptom.

We examined the extent to which the association between threshold-level depression (yes *vs* no) and dementia risk was accounted for by the identified symptoms. To this end, we estimated the percentage of attenuation of the association after adjusting effect estimates for each symptom individually and, in a final step, for all symptoms combined. A greater degree of attenuation indicates a larger proportion of the association explained by these symptoms.

We conducted a series of post-hoc sensitivity analyses. To test whether the observed symptom–dementia associations were robust to potential reverse causation due to prodromal or undiagnosed dementia, we repeated Cox proportional hazards regression analyses adjusted for age, sex, and ethnicity in a subsample of participants with at least 10 years of follow-up, excluding dementia cases that occurred within the first 10 years after baseline. To account for potential bias introduced due to missingness, models were repeated after additionally imputing missing data on depressive symptoms. We also estimated Fine–Gray competing risk models, treating death as a competing event for dementia, to test the robustness of results to informative censoring due to mortality. To address diagnostic uncertainty arising from inclusion of dementia diagnoses from the MHSDS, we re-estimated symptom–dementia associations after excluding dementia cases identified via this source. Additionally, we examined the cross-sectional associations of each dementia-related depressive symptom with established dementia risk factors, *APOEε4* status, and cognitive performance to better characterise the risk factor profiles of individuals presenting with these symptoms. We used tetrachoric correlations to examine the persistence of midlife depressive symptoms over time, conducted with measurements from baseline and a 10-year follow-up (2007–09). Higher correlation coefficients indicate greater stability of symptoms over time, with values higher than 0·4 interpreted as reflecting moderate persistence. A restriction analysis excluded individuals with any of the dementia-related symptoms to confirm that the association between midlife depression and dementia was attributable to the identified dementia-related depressive symptoms. To address the non-collapsibility of HRs that might have biased our attenuation analysis, we conducted a permutation (shuffling) analysis,[22] randomly reassigning any dementia-related symptoms across participants. Lastly, we calculated the age-specific dementia incidence rates in the

Whitehall II cohort to facilitate comparison with established population-based estimates.

Analyses were conducted with R Studio version 2024.04.2+764.

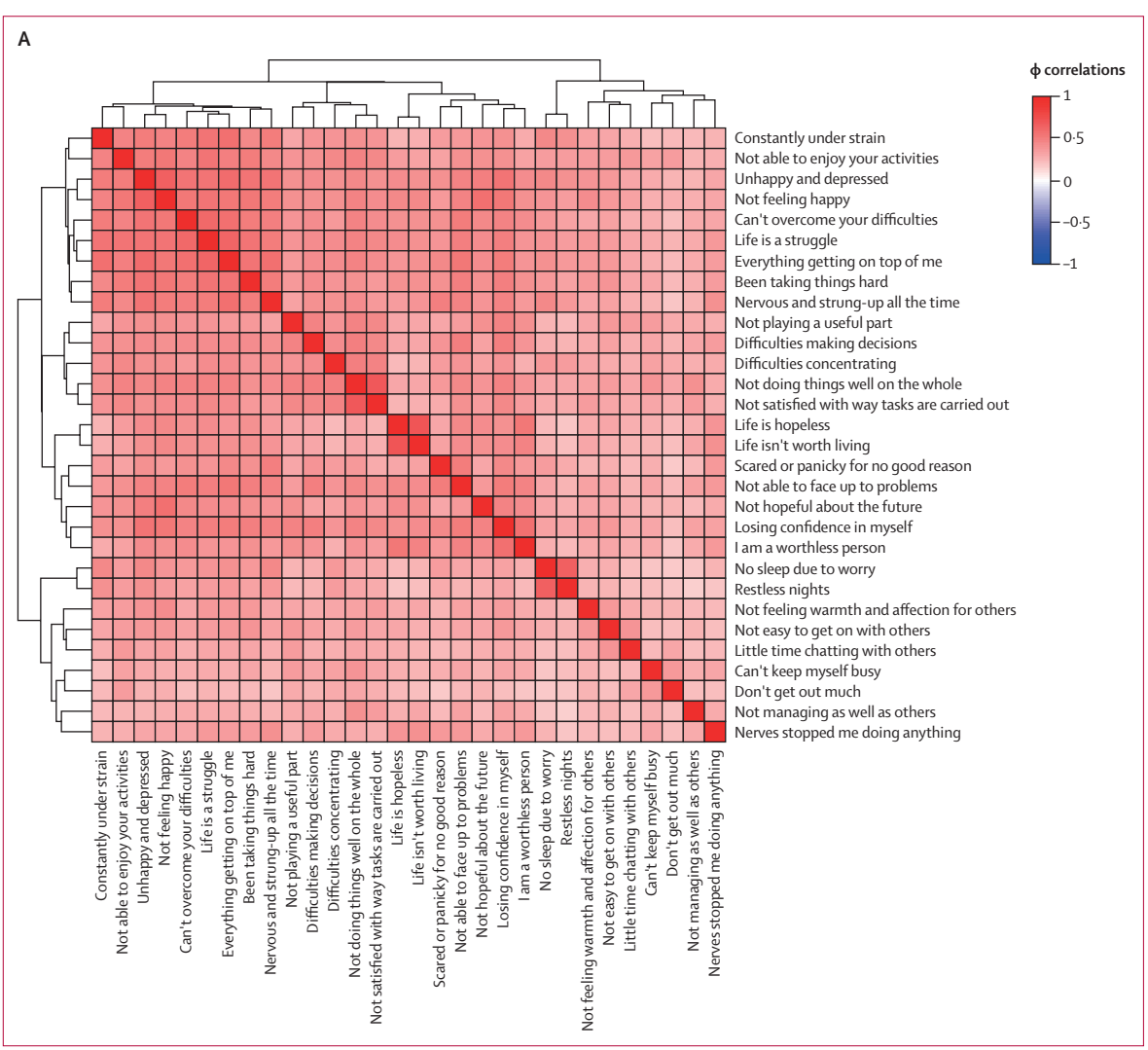
Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit for publication.

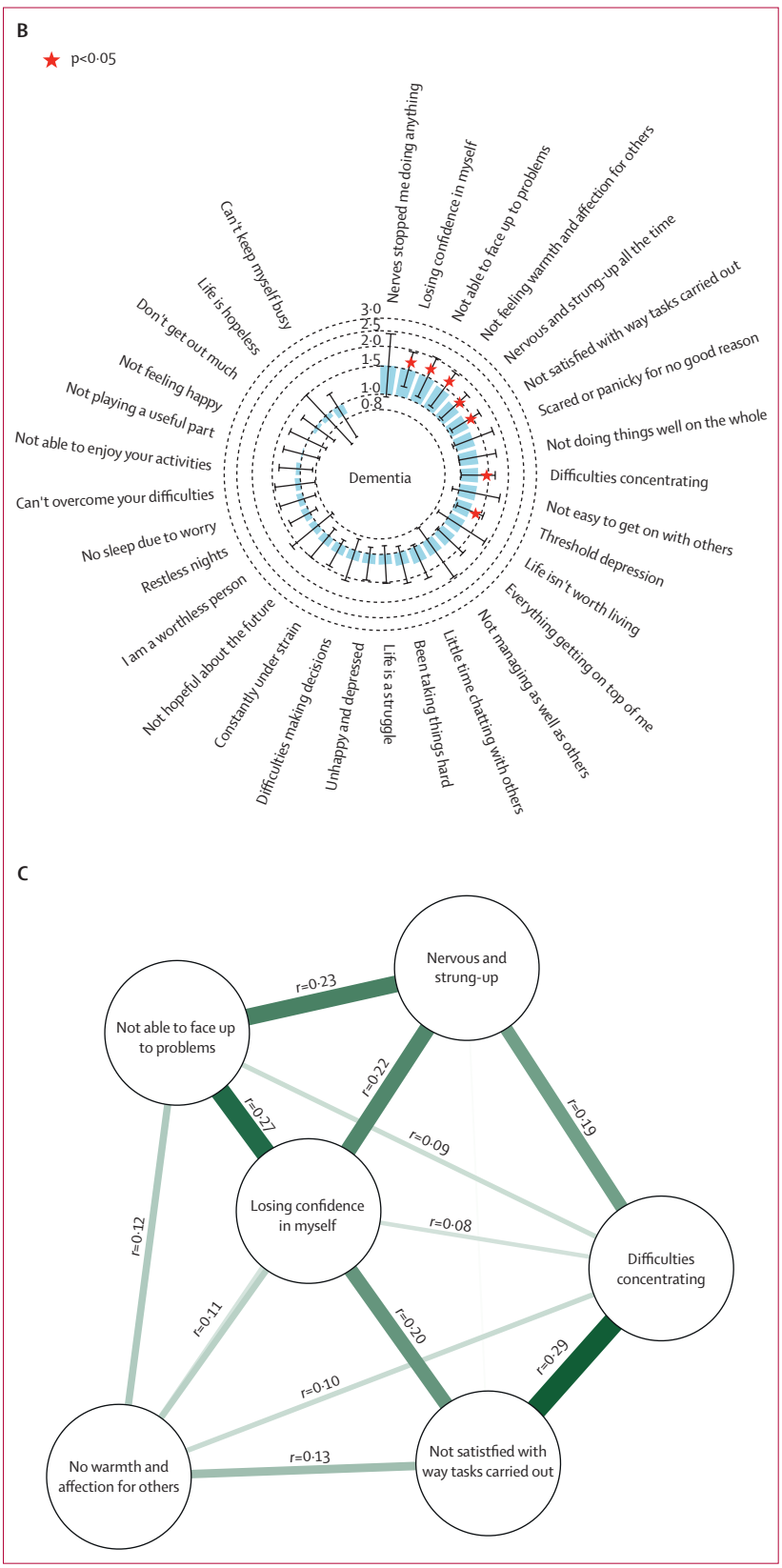
Results

Of the 6511 individuals who participated in a medical examination between April 24, 1997, and Jan 8, 1999 (the baseline for the present study), 740 were excluded (424 because of missing data on depression, 309 because of missing data on relevant clinical covariates, and seven because of missing linkage to electronic health records [appendix p 1]). The final analytical sample consisted of 5811 participants. The table shows the baseline characteristics of these participants, stratified by baseline depression status and dementia status at follow-up. The mean age of participants at baseline was 55·7 years (SD 6·0). 1646 (28·3%) participants were women, and 4165 (71·7%) were men. 5356 (92·2%) participants self-reported their ethnicity as White, and 455 (7·8%) reported their ethnicity as non-White. 1248 (21·5%) participants reported elevated depressive symptoms (GHQ score ≥5) at baseline, and 586 individuals (10·1%) developed dementia over the 23-year follow-up (mean follow-up 22·6 years [SD 5·0]). Age-specific dementia incidence increased gradually from midlife to older age, ranging from 0 cases per 1000 person-years before age 60 years to 39·0 cases per 1000 person-years at ages 85–89 years and 66·7 cases per 1000 person-years at ages 90 years and older (appendix p 4).

Compared with participants without elevated depressive symptoms, those with above-threshold-level



(Figure 1 continues on next page)



depression at baseline were more likely to be younger (p<0·0001), female (p<0·0001), obese (p=0·0055), and physically inactive (p<0·0001), and to have hearing impairment (p<0·0001), but were less likely to have hypertension (p<0·0001; table). Individuals who developed dementia were more likely to be older (p<0·0001), female (p=0·0027), from a non-White ethnic background (p=0·0006), and to have lower educational qualifications (p<0·0001) compared with those who remained dementia free; these individuals also had higher rates of diabetes (p<0·0001), hypertension (p<0·0001), elevated LDL cholesterol (p=0·0002), and hearing impairment (p=0·0006). Although *APOEε4* carriers were more likely than other participants to develop dementia over the follow-up period, they did not have an elevated risk of depression at baseline.

Among the 30 GHQ items, pairwise correlations ranged from 0·17 to 0·70, with clear clustering patterns emerging (figure 1A). The strongest associations (ϕ=0·70) were observed between "Life is hopeless" and "Life isn't worth living" and between "Not satisfied with the way tasks are carried out" and "Not being able to do things well on the whole". Strong correlations were also observed for symptoms related to sleep problems, such as "No sleep due to worry" and "Restless nights" (ϕ=0·66). Within-cluster correlations were also evident for "Unhappy and depressed", "Life is a struggle", and "Everything is getting on top of me" (ϕ≥0·55). Similarly, symptoms related to cognitive function, such as "Difficulties concentrating" and "Difficulties making decisions" clustered together (ϕ=0·43). Other moderately correlated clusters were symptoms related to self-esteem and hopelessness, including "Losing confidence in myself", "I am a worthless person", and "Not hopeful about the future" (ϕ≥0·42).

In analyses adjusted for age, sex, and ethnicity, six of the 30 GHQ symptoms were associated with long-term dementia risk (figure 1B): "Losing confidence in myself" (HR 1·51, 95% CI 1·16–1·96), "Not able to face up to problems" (1·49, 1·09–2·04), "Not feeling warmth and affection for others" (1·44, 1·06–1·95), "Nervous and strung-up all the time" (1·34, 1·03–1·72), "Not satisfied with the way tasks are carried out" (1·33, 1·05–1·69), and "Difficulties concentrating" (1·29, 1·01–1·65). Threshold-level depression was also associated with an increased risk of dementia (1·27, 1·03–1·56).

Symptom-specific associations were robust across several sensitivity analyses (appendix pp 5–7), including a lagged-onset analysis excluding dementia cases

Figure 1: **Cross-sectional correlations between depressive symptoms and prospective associations with incident dementia over a mean follow-up of 23 years**
(A) Clustering of depressive symptoms on the basis of pairwise correlations. (B) Association between depressive symptoms and dementia risk, adjusted for age, sex, and ethnicity. (C) Network structure of dementia-related depressive symptoms.

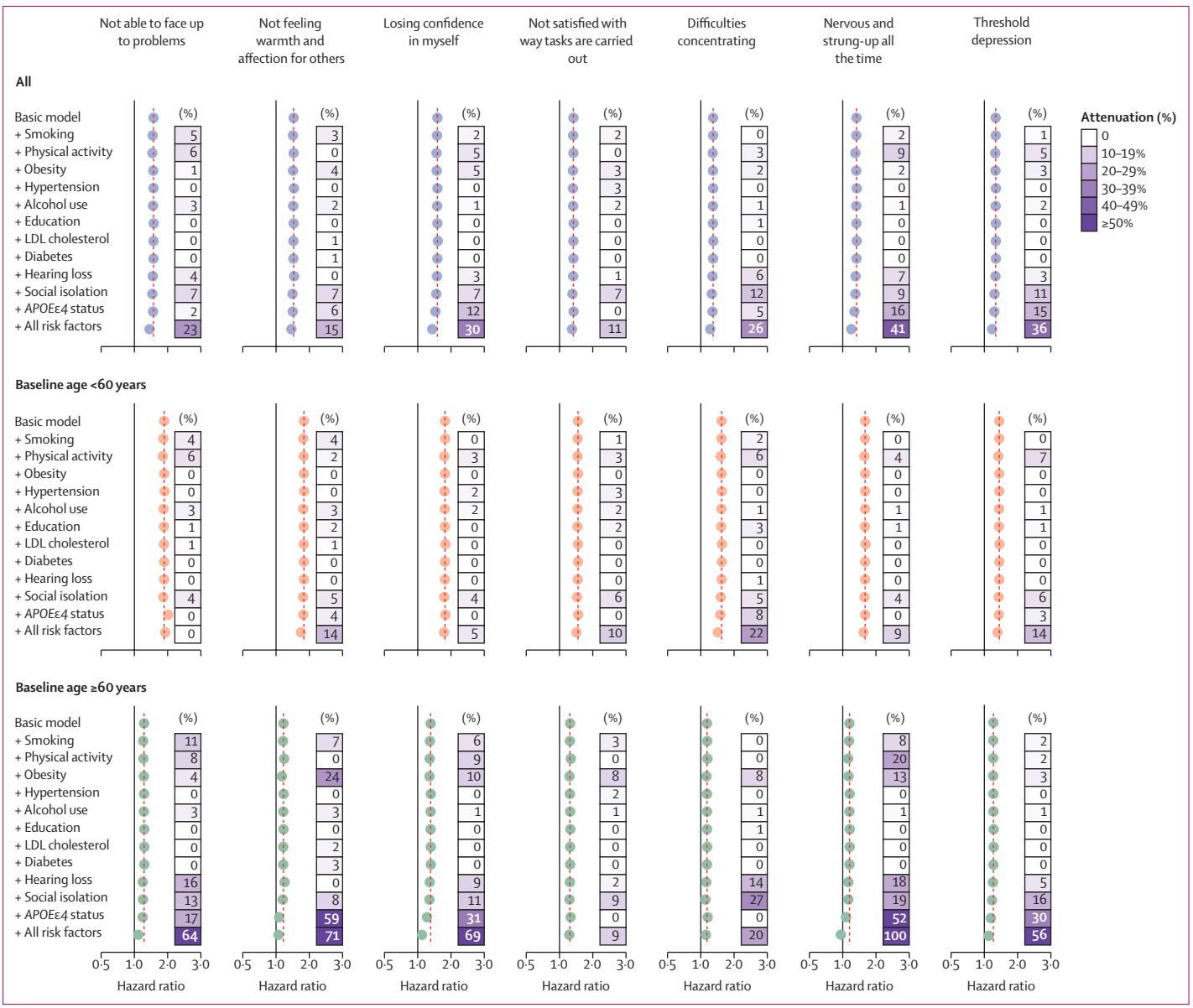


Figure 2: **Stratified and risk factor-adjusted prospective associations between depressive symptoms at baseline and dementia risk at follow-up of 23 years**
Associations were adjusted for 11 dementia risk factors and stratified by age. Dashed vertical lines indicate the hazard ratios of the basic model as a reference.

occurring within 10 years of baseline, analyses with imputed depression data, and Fine–Gray models accounting for death as a competing risk. In analyses excluding MHSDS-derived dementia cases, the number of incident cases decreased from 586 to 476; however, the associations (adjusted for age, sex, and ethnicity) between depressive symptoms and dementia remained materially unchanged.

In our regularised partial correlation network analysis to better understand the inter-relationships between the symptoms associated with long-term dementia risk, "Losing confidence in myself" emerged as a central node in the network, indicating that it was highly interconnected and might represent a core feature within the psychological distress profile linked to dementia risk (figure 1C).

In our analysis comparing participants younger than 60 years at baseline (n=3994; mean age 52·3 years [SD 3·7]) with those aged 60 years and older (n=1817; mean age 63·2 years [2·1]), we found that all symptom–dementia associations were stronger in individuals younger than 60 years (figure 2). In our examination of whether the associations between the six individual symptoms and dementia risk were influenced by the presence of established dementia risk factors, covering medical conditions, psychosocial factors, health

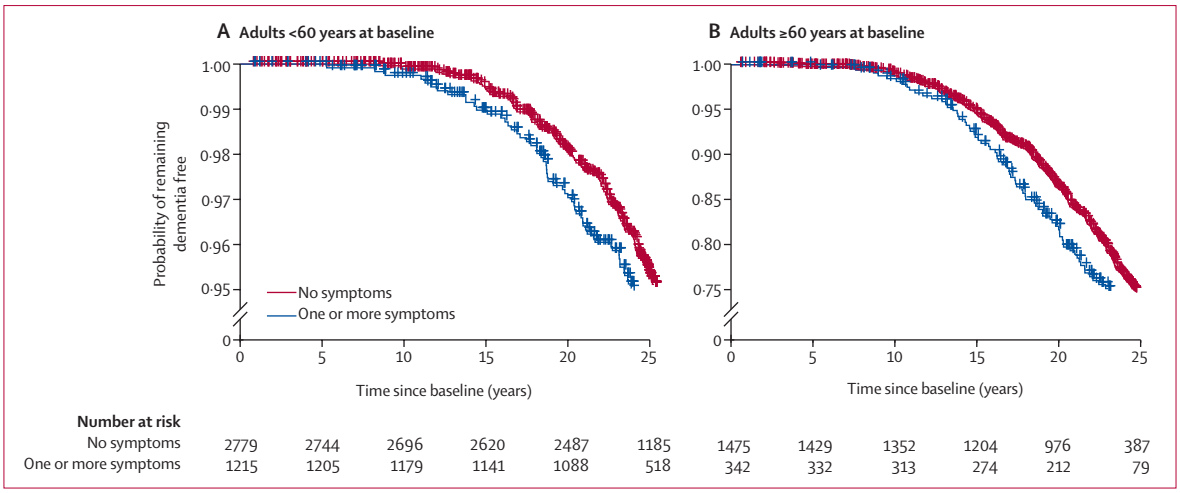


Figure 3: **Dementia-free probability by depressive symptom status and baseline age group**
Dementia-free probability by symptom status in middle-aged adults younger than 60 years (A) and adults aged 60 years and older (B) at baseline. Note that y-axes include axis breaks and that scales on y-axes differ between plots.

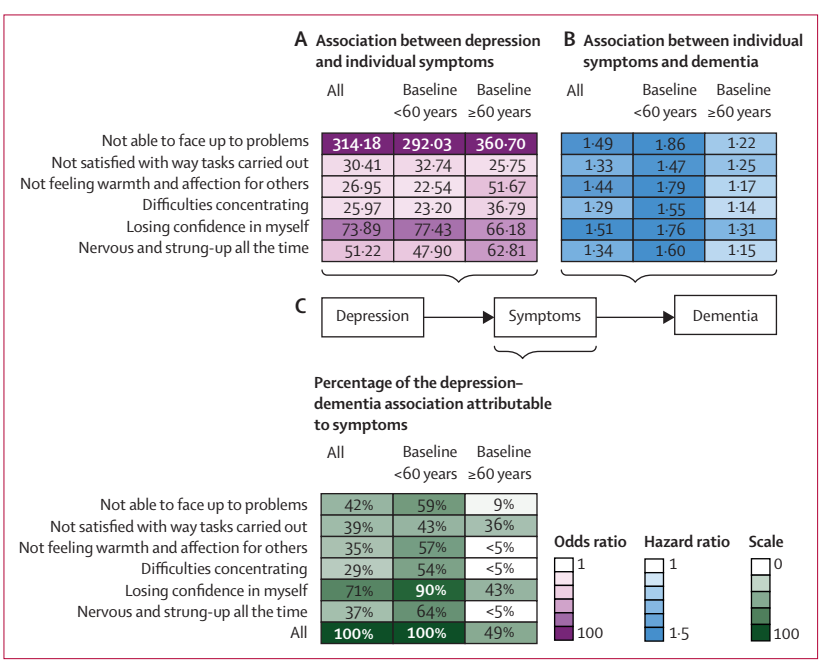


Figure 4: **Contribution of six dementia-related depressive symptoms to the dementia–depression association stratified by age**
(A) Odds ratios for the associations between depression and individual depressive symptoms. (B) Hazard ratios for the associations between individual depressive symptoms and dementia. (C) Percentage of depression–dementia association attributable to individual symptoms.

behaviours, and genetic risk, associations were also largely independent of the 11 other dementia risk factors, with the greatest attenuation of associations observed after adjustment for *APOEε4* status and hearing loss, particularly in older adults (figure 2, appendix pp 8–16).

In cross-sectional analyses, the six dementia-related depressive symptoms were most strongly associated with physical inactivity and hearing loss, as well as poorer memory and reasoning performance. These symptoms were also linked to higher alcohol consumption and, to a lesser extent, obesity (appendix p 17). Using measurements from baseline and the 10-year follow-up, we found evidence of moderate-to-strong persistence of the six symptoms over time. The tetrachoric correlation coefficient for the repeated assessments ranged from 0·43 to 0·45 for five of the six symptoms and was slightly higher for "I can't face up to problems" (r=0·53). Among symptoms not associated with increased dementia risk, the corresponding coefficients ranged from 0·32 to 0·58, with the lowest values for "I can't keep myself busy" and "I don't get out much" (both r=0·32) and the highest values for "Life is hopeless" (r=0·56) and "Life isn't worth living" (r=0·58; appendix p 18).

The effect of having one or more of the six symptoms at baseline on the probability of remaining dementia free in both younger (age <60 years at baseline) and older adults (age ≥60 years at baseline) is displayed visually in Kaplan–Meier plots; figure 3). Analyses of individual symptoms by age group confirmed a clear separation of curves across the entire follow-up period for all symptoms (appendix pp 19–20).

In our analyses of the associations between threshold-level depression and individual symptoms and between each symptom and incident dementia in subsamples of younger and older adults (figure 4A, B), all six symptoms were strongly associated with both threshold-level depression at baseline and incident dementia at follow-up, with the largest effect estimates observed for "Losing confidence in myself" and "Not able to face up to problems" across both analyses. In our estimation of the extent to which each symptom contributed to the overall depression–dementia association (figure 4C), when each individual symptom was adjusted for separately, "Losing confidence in myself" exerted the

greatest attenuating effect on the HR, accounting for 90% of the depression–dementia association in adults younger than 60 years and 43% in adults aged 60 years and older. Adjusting for all six symptoms at once fully attenuated the depression–dementia link in younger individuals and reduced the association by 49% in older individuals.

Additional analyses confirmed these results (appendix p 21). In exclusion analysis, the HR for threshold-level depression and dementia decreased from 1·27 to 0·77 after participants with any of the six symptoms were removed (n=4254; 431 dementia cases). In permutation analysis, the mean HR across all 1000 shuffled datasets for the link between threshold-level depression and dementia was 1·13, and none of the iterations produced an estimate as high as the HR observed in the main analysis.

Discussion

In our prospective cohort study of symptom-specific associations between midlife depression and dementia risk over a 22·6-year follow-up, using data from a large UK cohort, we found that the association between midlife depression and subsequent dementia was driven by a specific subset of symptoms: low self-confidence, cognitive and executive dysfunction, impaired social connection, and anxiety-related features. The observed symptom–dementia associations were largely independent of other established dementia risk factors, including *APOEε4* status, cardiometabolic conditions, and lifestyle-related risk factors. In individuals younger than 60 years at baseline, the association between midlife depression and dementia was fully explained by these symptoms.

Plausible mechanisms linking specific midlife depressive symptoms to an increased risk of dementia remain unclear. The six symptoms identified as most strongly associated with dementia were consistently related to established risk factors at baseline, including physical inactivity, hearing loss, and, to a lesser extent, excessive alcohol consumption. These findings suggest that behavioural and reserve-related mechanisms might partly contribute to the association of these six depressive symptoms with dementia. However, given that the symptom–dementia associations persisted after adjustment for these factors, these symptoms might also represent independent midlife risk factors for dementia. The associations of these symptoms with poorer cognition, particularly memory, raise the possibility that they might reflect early manifestations of mild cognitive impairment. This possibility is plausible, as amyloid pathology can precede the onset of clinical dementia by more than 15 years,[23] and the symptoms closely resemble the subjective complaints often made by people who are at the earliest stages of dementia. Other potential depression-related biological pathways that could underlie these associations include immune system disturbances, increased permeability of the blood–brain barrier,[24,25] and hypothalamic-pituitary-adrenal axis dysfunction.[8] These processes can, in turn, promote neuroinflammation, cerebrovascular damage, and hippocampal atrophy, thereby accelerating neurodegeneration and increasing vulnerability to Alzheimer's disease and other dementias of older age.[8] Notably, we found no consistent evidence that the associations between the six midlife symptoms and dementia were explained by genetic susceptibility (*APOEε4*), lower educational attainment, or adverse cardiometabolic risk factors such as high LDL cholesterol and diabetes.

Our findings align with previous studies reporting a modest association or no association between midlife depression, when measured as a single entity, and long-term dementia risk. For example, a study of 9745 Norwegian adults reported no difference in the prevalence of depression 22 years before dementia onset.[9] However, an association emerged when anxiety and depression symptoms were examined in combination, suggesting that co-occurring anxiety symptoms might increase dementia risk. Additionally, a 2022 meta-analysis of 33 studies[6] found an almost two-times increased risk of dementia for later-life depression, but it found no statistically significant association for early or midlife depression, and evidence in subgroup analyses of studies with follow-up durations of 15 years or more was inconsistent. The 2024 *Lancet* Commission[4] report on dementia prevention based its conclusion—to classify midlife depression as a dementia risk factor—on a subset of seven studies from the same meta-analysis. However, most of these studies had mean follow-up periods of less than 10 years and were focused on adults aged 60 years and older, with two of these studies[26,27] reporting mean baseline ages older than 70 years.

Interpretation of our findings requires consideration of various limitations. Our study was based on observational data, precluding the possibility of inferring causation. People with lived experience were not involved in the design of the study. Our depression measure was based on self-report and limited to 30 items, which might not fully capture the breadth and severity of depressive symptoms. Future research is warranted to examine symptom–dementia associations using more comprehensive, clinically validated assessments across multiple timepoints to gain further insights into symptom trajectories and the role of chronicity. Dementia ascertainment relied on national hospital and mortality records and the MHSDS, administrative sources that might miss milder or undiagnosed cases. Coding practices also vary: hospital and mortality records are based on specific ICD-10 diagnoses, whereas the MHSDS relies on less detailed coding. Moreover, hospital and mortality records often default to unspecified dementia in the absence of specialist assessments or biomarker evidence. Under-recognition is also common in specialist services, although diagnosis rates have increased over time. Given

the likely non-differential nature of these limitations between individuals with and without depressive symptoms, they are unlikely to have introduced major bias into our analyses. This interpretation is supported by the consistent results from sensitivity analyses restricted to ICD-10-diagnosed dementia cases and to those diagnosed during the more recent years of follow-up (ie, excluding dementia cases occurring within 10 years of baseline). It remains possible that the presence of depressive symptoms increases the likelihood of being diagnosed with dementia, contributing to an indication bias that could artificially inflate associations. However, such bias would mainly affect participants with threshold-level depression, who are more likely to access health-care services, rather than those with only one or a few specific symptoms associated with dementia, as seen in this study.

The generalisability of our findings should be interpreted with caution. Age-specific dementia incidence rates in the Whitehall II cohort were somewhat lower than those reported in general population cohorts, reflecting both the healthier occupational profile of civil servants and the under-ascertainment of dementia in linked electronic health records. The healthy worker effect likely leads to an underestimation of dementia incidence compared with the general population; however, studies comparing occupational and general population cohorts suggest that any resulting bias in risk factor–disease associations is unlikely to be substantial.[28] Moreover, our analytical sample included a lower proportion of women than men, limiting statistical power to examine sex differences in the associations between depressive symptoms and dementia. Further research is needed to establish whether our findings can be replicated in other study populations, including minority ethnic populations, and across different countries and settings.

Symptom-level approaches add value beyond traditional approaches that conceptualise depression as a single, uniform entity, obscuring clinically and biologically meaningful heterogeneity. Our findings, if replicated, could help clinicians distinguish between middle-aged patients whose depression reflects an elevated dementia risk and those whose symptoms are more likely due to other causes, supporting clinical evaluation and more tailored treatments.

Contributors

All authors participated in designing the study, generating hypotheses, interpreting the data, and critically reviewing the report. PF, with MK, had the primary responsibility for writing this paper. PF and JP conducted data analyses. All authors had access to the pseudonymised data reported in the study. PF, JP, and MK accessed and verified the data. All authors read and approved the final version of the manuscript and had final responsibility for the decision to submit for publication.

Declaration of interests

We declare no competing interests.

Data sharing

Whitehall II data are available for sharing within the scientific community. Bona fide researchers interested in accessing the data can apply through the Dementias Platform UK (https://www.dementiasplatform.uk) or the Whitehall Scientific committee (https://www.ucl.ac.uk/epidemiology-health-care/research/epidemiology-and-publichealth/research/whitehall-ii/data-sharing). Additional requests for clinical data from individual investigators can be submitted to the Whitehall II steering committees (whitehall2@ucl.ac.uk) and will be reviewed to ensure that data can be shared without compromising patient confidentiality. Participant-level electronic health records might be partly restricted on the basis of confidentiality or privacy laws and intellectual property restrictions.

Acknowledgments

This research was supported by the Wellcome Trust Collaborative Award in Science and the UK Medical Research Council infrastructure grant (Y014154). PF is supported by the Wellcome Trust (221854/Z/20/Z). JP is supported by the Research Council of Finland (350426). GDB is supported by the UK Medical Research Council (MR/P023444/1) and the US National Institute on Aging (1R56AG052519–01, 1R01AG052519–01A1). MK is supported by the Wellcome Trust (221854/Z/20/Z), the UK Medical Research Council (S011676, Y014154), the US National Institute on Aging (R01AG056477), and the Research Council of Finland (350426).

References

1 Nichols E, Steinmetz JD, Vollset SE, et al, and the GBD 2019 Dementia Forecasting Collaborators. 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* 2022; **7:** e105–25.

2 WHO. Global action plan on the public health response to dementia 2017–2025. World Health Organization, 2017.

3 Ferrari AJ, Santomauro DF, Aali A, et al, and the GBD 2021 Diseases and Injuries Collaborators. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* 2024; **403:** 2133–61.

4 Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the *Lancet* standing Commission. *Lancet* 2024; **404:** 572–628.

5 Leung DKY, Chan WC, Spector A, Wong GHY. Prevalence of depression, anxiety, and apathy symptoms across dementia stages: a systematic review and meta-analysis. *Int J Geriatr Psychiatry* 2021; **36:** 1330–44.

6 Stafford J, Chung WT, Sommerlad A, Kirkbride JB, Howard R. Psychiatric disorders and risk of subsequent dementia: systematic review and meta-analysis of longitudinal studies. *Int J Geriatr Psychiatry* 2022; **37:** gps.5711.

7 Singh-Manoux A, Dugravot A, Fournier A, et al. Trajectories of depressive symptoms before diagnosis of dementia: a 28-year follow-up study. *JAMA Psychiatry* 2017; **74:** 712–18.

8 Huang Y-Y, Gan Y-H, Yang L, Cheng W, Yu J-T. Depression in Alzheimer's disease: epidemiology, mechanisms, and treatment. *Biol Psychiatry* 2024; **95:** 992–1005.

9 Aunsmo RH, Strand BH, Anstey KJ, et al. Associations between depression and anxiety in midlife and dementia more than 30 years later: the HUNT study. *Alzheimers Dement (Amst)* 2024; **16:** e70036.

10 Fried EI. The 52 symptoms of major depression: lack of content overlap among seven common depression scales. *J Affect Disord* 2017; **208:** 191–97.

11 Fried EI, Flake JK, Robinaugh DJ. Revisiting the theoretical and methodological foundations of depression measurement. *Nat Rev Psychol* 2022; **1:** 358–68.

12 Feczko E, Miranda-Dominguez O, Marr M, Graham AM, Nigg JT, Fair DA. The heterogeneity problem: approaches to identify psychiatric subtypes. *Trends Cogn Sci* 2019; **23:** 584–601.

13 Marmot M, Brunner E. Cohort profile: the Whitehall II study. *Int J Epidemiol* 2005; **34:** 251–56.

14 Goldberg DP. The detection of psychiatric illness by questionnaire. Maudsley monograph no. 2. Oxford University Press, 1972.

15 Stansfeld SA, Marmot MG. Social class and minor psychiatric disorder in British civil servants: a validated screening survey using the General Health Questionnaire. *Psychol Med* 1992; **22:** 739–49.

16 Head J, Stansfeld SA, Ebmeier KP, et al. Use of self-administered instruments to assess psychiatric disorders in older people: validity of the General Health Questionnaire, the Center for Epidemiologic Studies Depression Scale and the self-completion version of the revised Clinical Interview Schedule. *Psychol Med* 2013; **43:** 2649–56.
17 Kivimäki M, Singh-Manoux A, Batty GD, et al. Association of alcohol-induced loss of consciousness and overall alcohol consumption with risk for dementia. *JAMA Netw Open* 2020; **3:** e2016084.
18 WHO. Global recommendations on physical activity for health. World Health Organization, 2010.
19 Rusmaully J, Dugravot A, Moatti J-P, et al. Contribution of cognitive performance and cognitive decline to associations between socioeconomic factors and dementia: a cohort study. *PLoS Med* 2017; **14:** e1002334.
20 Sommerlad A, Perera G, Singh-Manoux A, Lewis G, Stewart R, Livingston G. Accuracy of general hospital dementia diagnoses in England: sensitivity, specificity, and predictors of diagnostic accuracy 2008-2016. *Alzheimers Dement* 2018; **14:** 933–43.
21 van Bueren S. Flexible imputation of missing data, 2nd edn. Chapman and Hall/CRC, 2018.
22 Breiman L. Random forests. *Mach Learn* 2001; **45:** 5–32.
23 Hampel H, Hardy J, Blennow K, et al. The amyloid-β pathway in Alzheimer's disease. *Mol Psychiatry* 2021; **26:** 5481–503.
24 Hayley S, Hakim AM, Albert PR. Depression, dementia and immune dysregulation. *Brain* 2021; **144:** 746–60.
25 Frank P, Jokela M, Batty GD, Cadar D, Steptoe A, Kivimäki M. Association between systemic inflammation and individual symptoms of depression: a pooled analysis of 15 population-based cohort studies. *Am J Psychiatry* 2021; **178:** 1107–18.
26 Mirza SS, de Bruijn RF, Direk N, et al. Depressive symptoms predict incident dementia during short- but not long-term follow-up period. *Alzheimers Dement* 2014; **10** (suppl)**:** S323–29.
27 Richard E, Reitz C, Honig LH, et al. Late-life depression, mild cognitive impairment, and dementia. *JAMA Neurol* 2013; **70:** 374–82.
28 Batty GD, Shipley M, Tabák A, et al. Generalizability of occupational cohort study findings. *Epidemiology* 2014; **25:** 932–33.