



Article

<https://doi.org/10.1038/s41562-024-02078-1>

Plasma proteomic signatures of social isolation and loneliness associated with morbidity and mortality

Received: 17 January 2024

Chun Shen^{1,2,3}, Ruohan Zhang⁴, Jintai Yu⁵, Barbara J. Sahakian^{1,6,7}✉, Wei Cheng^{1,2,5}✉ & Jianfeng Feng^{1,2,4,8,9}✉

Accepted: 31 October 2024

Published online: 03 January 2025

Check for updates

The biology underlying the connection between social relationships and health is largely unknown. Here, leveraging data from 42,062 participants across 2,920 plasma proteins in the UK Biobank, we characterized the proteomic signatures of social isolation and loneliness through proteome-wide association study and protein co-expression network analysis. Proteins linked to these constructs were implicated in inflammation, antiviral responses and complement systems. More than half of these proteins were prospectively linked to cardiovascular disease, type 2 diabetes, stroke and mortality during a 14 year follow-up. Moreover, Mendelian randomization (MR) analysis suggested causal relationships from loneliness to five proteins, with two proteins (ADM and ASGR1) further supported by colocalization. These MR-identified proteins (GFRA1, ADM, FABP4, TNFRSF10A and ASGR1) exhibited broad associations with other blood biomarkers, as well as volumes in brain regions involved in interoception and emotional and social processes. Finally, the MR-identified proteins partly mediated the relationship between loneliness and cardiovascular diseases, stroke and mortality. The exploration of the peripheral physiology through which social relationships influence morbidity and mortality is timely and has potential implications for public health.

Social relationships are adaptive and critical for wellbeing and survival in social species¹. Social isolation and loneliness, characterized as reflections of objective and subjective manifestations of impoverished social relationships, are increasingly recognized as important global public concerns². Cumulative evidence demonstrates that both social isolation and loneliness are linked to morbidity and mortality, with effects comparable to traditional risk factors such as smoking and obesity^{3–6}. Despite these empirical associations, the

underlying mechanisms through which social relationships impact health remain elusive.

Experimental studies show that social interactions can causally alter animal physiology, such as sympathetic nervous system (SNS) and hypothalamic–pituitary–adrenal (HPA) activity, inflammation and antiviral responses, and directly influence disease risk^{7–9}. These patterns parallel observations in human correlational studies^{10,11}. Moreover, adverse social relationships have been associated with unhealthy

¹Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China. ²Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, Shanghai, China. ³Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK. ⁴Department of Computer Science, University of Warwick, Coventry, UK. ⁵Department of Neurology and National Center for Neurological Disorders, Huashan Hospital, State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Fudan University, Shanghai, China. ⁶Department of Psychiatry, University of Cambridge, Cambridge, UK. ⁷Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK. ⁸Zhangjiang Fudan International Innovation Center, Shanghai, China. ⁹School of Data Science, Fudan University, Shanghai, China. ✉e-mail: bjs1001@cam.ac.uk; wcheng@fudan.edu.cn; jffeng@fudan.edu.cn

Table 1 | Characteristics of study population at baseline in the UK Biobank

	All (N=42,062)	Social isolation at baseline			Loneliness at baseline		
		Not isolated (N=38,157)	Socially isolated (N=3,905)	P value	Not lonely (N=39,373)	Lonely (N=2,689)	P value
Age at baseline, in years	56.4 (8.2)	56.4 (8.2)	56.7 (7.9)	0.014	56.4 (8.2)	56.2 (8.1)	0.140
Sex				0.426			0.812
Female	21,986 (52.3%)	19,969 (52.3%)	2,017 (51.7%)		20,574 (52.3%)	1,412 (52.5%)	
Male	20,076 (47.7%)	18,188 (47.7%)	1,888 (48.3%)		18,799 (47.7%)	1,277 (47.5%)	
Ethnicity				2.9×10^{-7}			0.001
White	38,274 (91.0%)	34,814 (91.2%)	3,460 (88.6%)		35,887 (91.1%)	2,387 (88.8%)	
Mixed	1,718 (4.1%)	1,535 (4.0%)	183 (4.7%)		1,580 (4.0%)	138 (5.1%)	
Asian	1,503 (3.6%)	1,320 (3.5%)	183 (4.7%)		1,387 (3.5%)	116 (4.3%)	
Black	188 (0.4%)	162 (0.4%)	26 (0.7%)		174 (0.4%)	14 (0.5%)	
Other	379 (0.9%)	326 (0.9%)	53 (1.4%)		345 (0.9%)	34 (1.3%)	
Education level				$<2.2 \times 10^{-16}$			$<2.2 \times 10^{-16}$
College/university degree	14,816 (35.2%)	13,574 (35.6%)	1,242 (31.8%)		14,129 (35.9%)	687 (25.5%)	
Education to age 18 years	14,083 (33.5%)	12,878 (33.8%)	1,205 (30.8%)		13,169 (33.4%)	914 (34.0%)	
Education to age 16 years	6,863 (16.3%)	6,271 (16.4%)	592 (15.2%)		6,372 (16.2%)	491 (18.3%)	
No qualifications	6,300 (15.0%)	5,434 (14.2%)	866 (22.2%)		5,703 (14.5%)	597 (22.2%)	
Household income				$<2.2 \times 10^{-16}$			$<2.2 \times 10^{-16}$
At least £31,000	21,334 (50.7%)	20,107 (52.7%)	1,227 (31.4%)		20,392 (51.8%)	942 (35.0%)	
Less than £31,000	20,728 (49.3%)	18,050 (47.3%)	2,678 (68.6%)		18,981 (48.2%)	1,747 (65.0%)	
Current smoker				$<2.2 \times 10^{-16}$			$<2.2 \times 10^{-16}$
Yes	4,450 (10.6%)	3,721 (9.8%)	729 (18.7%)		3,974 (10.1%)	476 (17.7%)	
No	37,612 (89.4%)	34,436 (90.2%)	3,176 (81.3%)		35,399 (89.9%)	2,213 (82.3%)	
Alcohol intake				$<2.2 \times 10^{-16}$			$<2.2 \times 10^{-16}$
At least three times per week	18,868 (44.9%)	17,556 (46.0%)	1,312 (33.6%)		17,886 (45.4%)	982 (36.5%)	
Twice or less per week	23,194 (55.1%)	20,601 (54.0%)	2,593 (66.4%)		21,487 (54.6%)	1,707 (63.5%)	
Body mass index, kg m⁻²	27.4 (4.8)	27.4 (4.7)	27.9 (5.3)	6.0×10^{-8}	27.4 (4.7)	28.5 (5.5)	$<2.2 \times 10^{-16}$

Two-sided P values were estimated by a t-test for continuous variables and chi-squared test for categorical variables.

lifestyles¹², potentially impacting these physiological pathways and subsequently affecting health. In comparison with behavioural moderators that indirectly influence health, there is a growing focus on understanding biological processes mediating the link between social relationships and health, given their relevance to improving disease prediction, prevention and intervention. It is noteworthy that proteins, as the final products of gene expression, serve as the main functional components of biological processes and represent a major source of drug targets¹³. Therefore, understanding the proteomic associations of social isolation and loneliness becomes imperative for unravelling the biology underpinning the effects of social relationships on health. One study reported that circulating brain-derived neurotrophic factor levels were associated with social relationships and partly mediated the effect between social support and dementia risk¹⁴. However, no comprehensive proteome-wide association study (PWAS) of social isolation and loneliness has been performed so far.

Here, leveraging high-throughput, population-scale proteomics alongside deep phenotypic data from the UK Biobank, we aimed, in this novel study, to answer two key questions: (1) What are the proteomic profiles associated with social isolation and loneliness? (2) How do proteomic alterations contribute to the relationship of social isolation, loneliness and health? To address the first question, we initially conducted PWASs and protein co-expression network analysis for social isolation and loneliness, respectively. The identified proteins and protein modules were subsequently examined potential causal

relationships with social isolation and loneliness using bidirectional Mendelian randomization (MR) and colocalization analyses. To explore the relationship between the MR-identified proteins and broad physical functions, we investigated their associations with other blood biomarkers. On the basis of the social brain hypothesis¹⁵ and increasing research on the neurobiology of social isolation and loneliness¹⁶, we further related these proteins to brain volumes. For the second question, we delved into the prospective associations between proteins linked to social isolation and loneliness and the incidence of morbidity and mortality. Specifically, we focused on five diseases with well-documented associations with social relationships: cardiovascular diseases (CVDs)^{17,18}, dementia^{19,20}, type 2 diabetes (T2D)²¹, depression²² and stroke^{18,23}. Finally, we evaluated the role of plasma proteins in the connection between social relationships and the risk of morbidity and mortality in mediation analyses for time-to-event outcomes.

Results

Cohort characteristics

Our primary study population included 42,062 participants (aged 56.4 ± 8.2 years, 52.3% female) from the UK Biobank who had quality-controlled proteomic data and complete behavioural data including social isolation, loneliness and all covariates. A flow chart of the participant selection is shown in Supplementary Fig. 1. Among these, 3,905 (9.3%) reported being socially isolated and 2,689 (6.4%) felt lonely. Detailed demographic characteristics, stratified by social

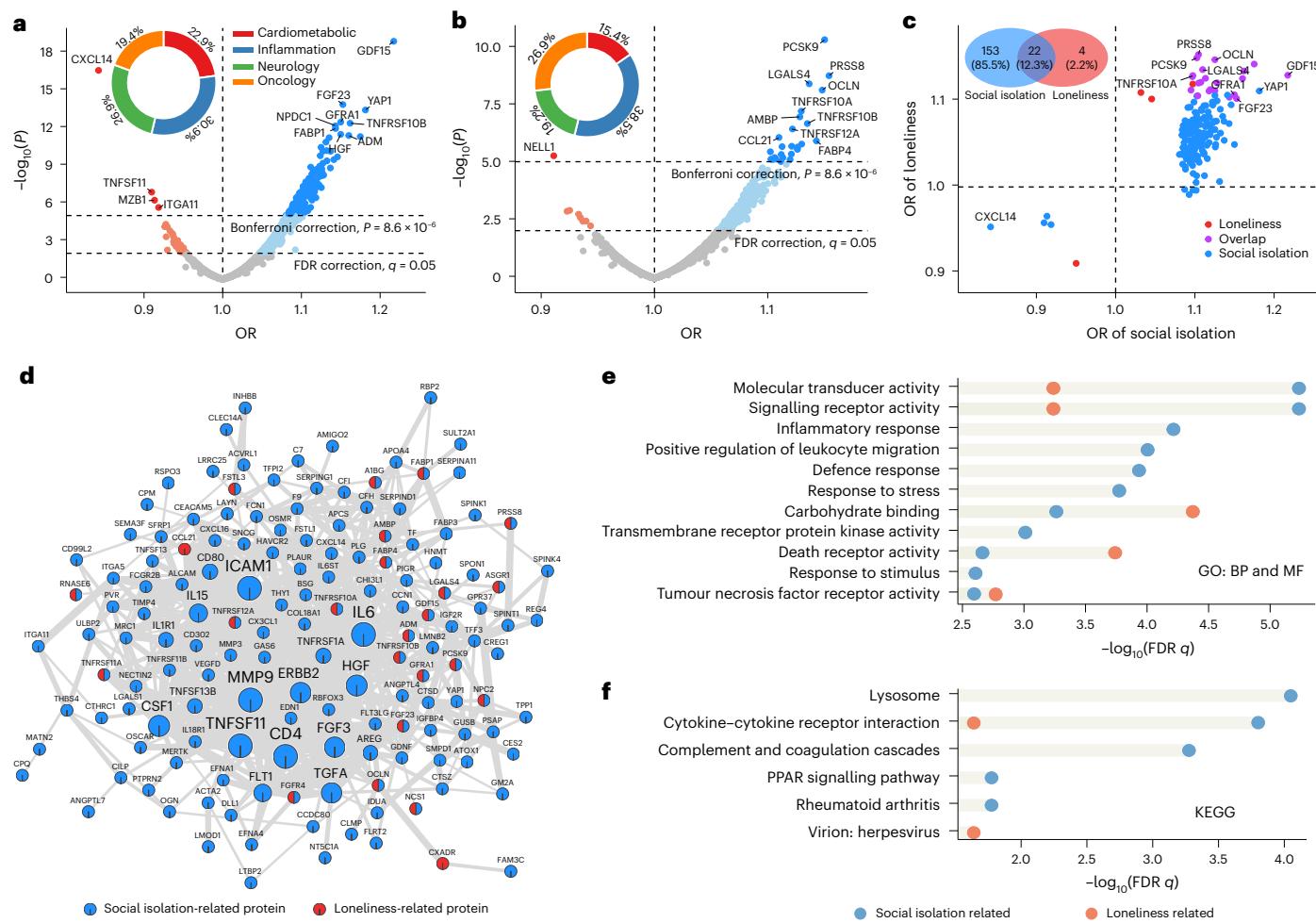


Fig.1 | PWAS for social isolation and loneliness. The average sample size for the 2,920 proteins analysed is 37,704, ranging from 30,778 to 41,396. Sample sizes for proteins significantly associated with social isolation and loneliness after Bonferroni correction ($P < 0.05 / (2,920 \times 2) = 8.6 \times 10^{-6}$) are detailed in Supplementary Tables 1 and 2. Logistic regression models were adjusted for age, sex, site, batch, time gap between blood collection and protein measurement, ethnicity, education level, household income, smoking, alcohol consumption, BMI and the first 20 genetic PCs. All statistical tests were two sided. **a**, A volcano plot displaying the correlation of protein abundance with social isolation. The x axis represents ORs, and the y axis represents $-\log_{10}(P\text{values})$. The dashed lines indicate the thresholds for Bonferroni and FDR corrections ($q < 0.05$) when considering social isolation and loneliness simultaneously. The pie charts depict the proportions of identified proteins across four distinct panels.

isolation and loneliness, are presented in Table 1. During a median (s.d.) follow-up of 13.7 (2.1) years (ending on 30 November 2022 or death), 2,695 participants developed CVD, 892 developed all-cause dementia, 1,703 developed T2D, 1,521 developed depression, 983 developed stroke and 4,255 passed away.

Proteins associated with social isolation and loneliness

We conducted logistic regression for PWAs involving 2,920 plasma proteins, using social isolation or loneliness as the outcome. In simple models incorporating age, sex, site, technical factors and the first 20 genetic principal components (PCs) as covariates, we found 776 proteins significantly associated with social isolation and 519 proteins associated with loneliness ($P < 0.05 / (2,920 \times 2) = 8.6 \times 10^{-6}$) (Supplementary Fig. 2). After additional adjustments for ethnicity, education level, household income, smoking, alcohol consumption and body mass index (BMI), 175 proteins associated with social isolation (Fig. 1a

and Supplementary Table 1) and 26 proteins associated with loneliness (Fig. 1b and Supplementary Table 2) maintained significance at the Bonferroni-corrected threshold.

Notably, growth differentiation factor 15 (GDF15), a protein belonging to the transforming growth factor- β superfamily that acts as an inflammatory marker²⁴, demonstrated the strongest association with social isolation (odds ratio (OR) of 1.22, 95% confidence interval (CI) 1.17 to 1.27, $P = 1.2 \times 10^{-19}$, $N = 41,396$), while proprotein convertase subtilisin/kexin type 9 (PCSK9), a key protein in the regulation of cholesterol metabolism²⁵, showed the strongest association with loneliness (OR of 1.15, 95% CI 1.10 to 1.20, $P = 4.2 \times 10^{-11}$, $N = 41,396$). The majority of identified proteins exhibited a positive association, indicating that higher protein abundance was linked to an elevated risk of social isolation or loneliness. However, only four proteins emerged as protective factors against social isolation, and one against loneliness. Particularly notable among these was C-X-C motif chemokine ligand-14 (CXCL14),

an immune and inflammatory modulator²⁶, emerging as the second most significant protein associated with social isolation (OR of 0.84, 95% CI 0.81 to 0.88, $P = 2.4 \times 10^{-17}$, $N = 40,090$).

Furthermore, 22 proteins (12.3% of the overall identified proteins) were common (Fig. 1c), and the proteome-wide associative patterns of social isolation and loneliness were moderately related ($r = 0.54$, 95% CI 0.52 to 0.57, $P < 2.2 \times 10^{-16}$, $N = 2,920$; Supplementary Fig. 3), implicating shared and distinct proteomic signatures between social isolation and loneliness. Restricted cubic splines²⁷ revealed that four proteins displayed a significant nonlinear association with social isolation ($P < 0.05/(175 + 26) = 2.5 \times 10^{-4}$; Supplementary Fig. 4), while no nonlinear proteomic association with loneliness was detected.

Sensitivity analyses

First, ordered logistic regressions using raw scores of social isolation and loneliness were performed to test the robustness and potential dose-dependent associations with plasma protein levels. This approach yielded results consistent with those from dichotomous variables and logistic regression but with greater statistical power (580 and 125 proteins associated with social isolation and loneliness, respectively; Supplementary Fig. 5).

Given the reported sampling bias in the UK Biobank Pharma Proteomics Project (UKB-PPP) subcohort composition, we replicated the primary analyses specifically within the randomly selected subset ($N = 36,250$), which is highly representative of the overall UK Biobank population. The proteins identified in this subset were consistent with those in all available samples and the proteomic associative patterns in these two populations were highly correlated (both $r > 0.9$, $P < 2.2 \times 10^{-16}$; Supplementary Fig. 6). Additionally, to mitigate potential population stratification, we conducted PWAS specifically in Caucasians ($N = 35,697$) and observed that the proteomic associative patterns were highly correlated with those found in the full sample (both $r > 0.9$, $P < 2.2 \times 10^{-16}$; Supplementary Fig. 7).

Next, interaction terms between sex or age and protein levels were included in logistic models to assess potential sex and age differences. No significant interaction effects between sex or age and proteins associated with social isolation or loneliness were observed ($P > 0.05/(175 + 26) = 2.5 \times 10^{-4}$). The PWAS results for males ($N = 20,076$) and females ($N = 21,986$) are shown in Supplementary Fig. 8. Additionally, the PWAS results for younger (<60 years, $N = 24,171$) and older (≥ 60 years, $N = 17,891$) groups are shown in Supplementary Fig. 9.

Considering the established link between loneliness and depression²², we explored the potential influence of depressive symptoms on the proteomic association of social isolation and loneliness ($N = 38,778$). The adjustment for depressive symptoms modestly impacted the association between social isolation and proteins (Supplementary Fig. 10a,b), whereas notably weakening the association between loneliness and proteins (for example, the OR of GDF15 decreased by 9.5% after adjustment; Supplementary Fig. 10c–e). Additionally, we examined the potential influence of physical activity ($N = 34,548$) and found that the associations between social isolation, loneliness and plasma proteins were largely independent of physical activity levels (Supplementary Fig. 11).

Incorporating both social isolation and loneliness into the model had minimal impact on their respective associations with proteins (Supplementary Fig. 12). To further investigate the underlying construct of impoverished social relationships, multinomial logistic regressions were conducted to test the association between the four-group classification and plasma protein levels. Compared with participants who were neither isolated nor lonely ($N = 36,100$), 116 proteins differed significantly in the socially isolated (SI) but not lonely (LO) group (SI+LO−; $N = 3,273$), 8 in the not isolated but lonely group (SI−LO+; $N = 2,057$) and 22 in the socially isolated and lonely group (SI+LO+; $N = 632$) (Supplementary Fig. 13). Interestingly, GDF15 was identified

as the top differentiated protein in SI+LO− and SI+LO+, while PCSK9 was the top differentiated protein in SI−LO+.

Finally, we performed cross-validation by randomly splitting the UK Biobank samples 100 times. Our results showed that most of the significant proteins identified using the full sample retained significance in at least one of the two split samples, and proteomic associative patterns for social isolation (mean $r = 0.56$, s.d. 0.05) and loneliness (mean $r = 0.35$, s.d. 0.08) between the two split samples exhibited medium to large correlations (Supplementary Fig. 14).

PPI and pathway enrichment

To explore potential interactions among proteins associated with social isolation or loneliness, we analysed the protein–protein interaction (PPI) networks within the identified pool of 179 proteins using the STRING database²⁸. A prominent interconnected network featuring 150 proteins and 690 PPIs was found (Fig. 1d). The maximal clique centrality (MCC) method²⁹ revealed the top three hub proteins of interleukin 6 (IL6), intercellular adhesion molecule-1 (ICAM1) and cluster of differentiation 4 (CD4) all exhibiting associations with social isolation.

Furthermore, we investigated functional pathways for the proteins associated with social isolation and loneliness at a false discovery rate (FDR) $q < 0.05$. We observed several shared pathways between social isolation and loneliness. Both sets of proteins exhibited significant enrichment in immune pathways such as cytokine–cytokine receptor interaction and antiviral processes (Fig. 1e,f and Supplementary Tables 3 and 4). Additionally, proteins associated with social isolation specifically showed enrichment in complement system and mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase (ERK) signalling.

Protein networks linked to social isolation and loneliness

Given that complex diseases are not caused by individual proteins, but rather result from highly interactive protein networks³⁰, we applied a data-driven approach to classify plasma proteins into clusters or modules based on protein co-expression patterns. Thirteen non-overlapping protein modules were identified, ranging in size between 30 and 1,051 proteins (Supplementary Fig. 15). We then correlated the module eigengene with social isolation and loneliness, revealing that M4, M8 and M12 showed associations with both conditions after Bonferroni correction ($P < 0.05/(13 \times 2) = 0.002$) (Fig. 2a,b). Additionally, M3 was found to be associated with social isolation. Enrichment analyses suggested distinct functions associated with these modules. For example, protein module M4 exhibited enrichment in immune-related pathways (Fig. 2c) and included the top protein related to social isolation, GDF15 (Supplementary Table 5). M8 demonstrated enrichment in metabolic processes (Fig. 2d and Supplementary Table 6). M12 showed enrichment in neutrophil degranulation (Fig. 2e and Supplementary Table 7). M3 was enriched in complement systems (Fig. 2f) and included the top protein related to loneliness, PCSK9 (Supplementary Table 8). Actually, M3 was also associated with loneliness if relaxing the significance threshold to FDR correction. Sensitivity analysis demonstrated that the biological relevance of the identified modules associated with social isolation and loneliness remained robust across various co-expression network construction parameters (Supplementary Figs. 16–18).

Moreover, we examined the relationship between significant modules with the proteins identified by PWAS using two-sided Fisher's exact tests. After Bonferroni correction, both protein sets associated with social isolation and loneliness were significantly enriched in M4, and the results were consistent across all proteins, the top 20% and the top 10% of proteins in each module (Supplementary Table 9). It is noteworthy that both approaches—the PWAS and the protein co-expression networks—were consistent and complementary, together providing more comprehensive information.

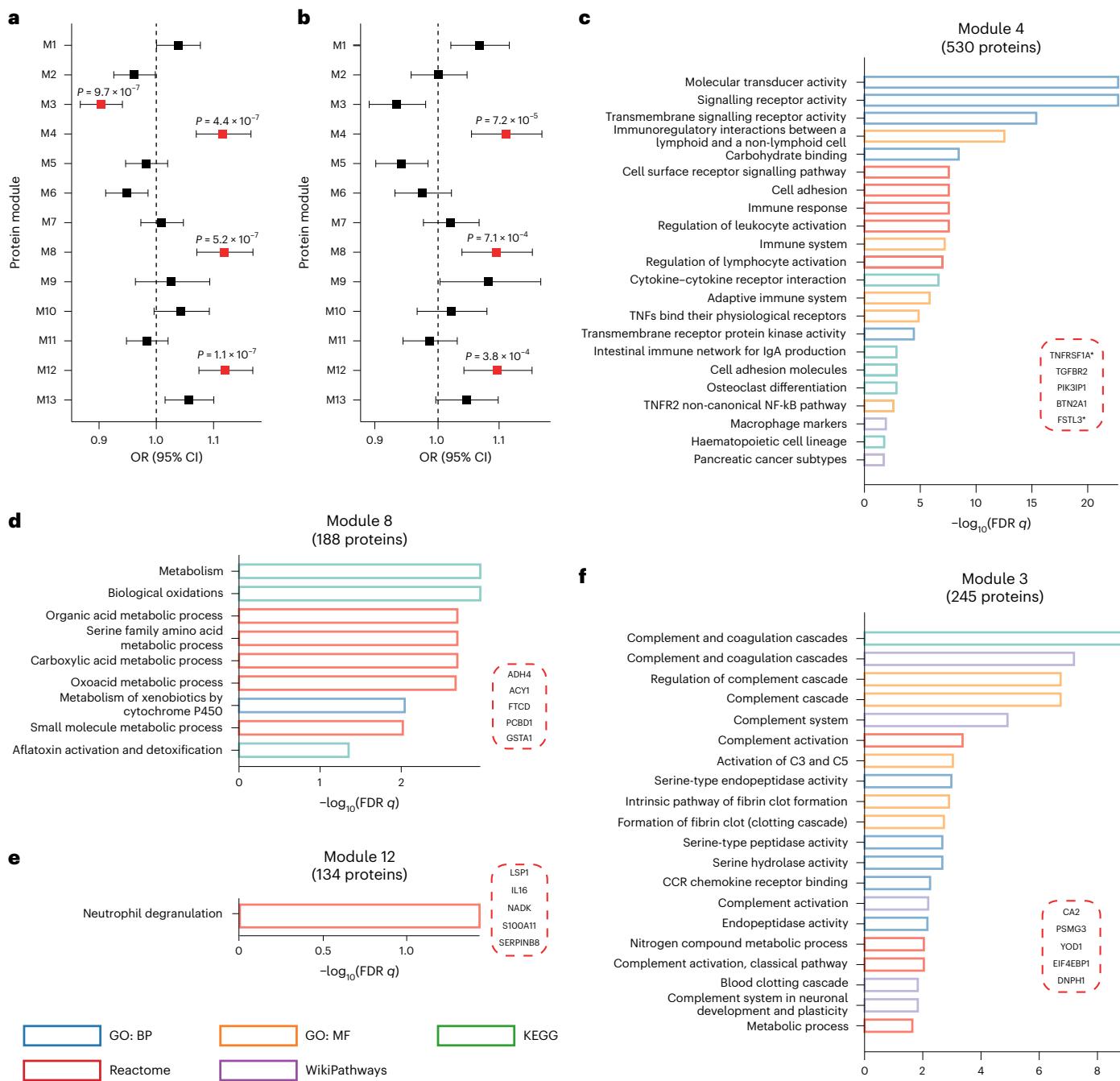


Fig. 2 | Association of protein co-expression network with social isolation and loneliness ($N = 35,475$). a, b, A forest plot showing the associations between protein modules and social isolation. M1 to M13 represent 13 modules identified through protein network analysis. Two-sided P values were derived from logistic regression models adjusted for the same covariates as in the PWAS. The dots represent ORs, and error bars indicate 95% CIs. The red markers denote significance after Bonferroni correction ($P < 0.05/26 = 0.002$). **c–f,** Enrichment analysis of proteins from modules 4 (c), 8 (d), 12 (e) and 3 (f). The top five significantly enriched pathways from each database are displayed, with the x-axis representing $-\log_{10}(\text{FDR } q)$ values. The top five proteins in each module that are most highly correlated with the overall network expression are displayed in the dashed box. Red indicates a positive correlation, while blue indicates a negative correlation. Proteins marked with asterisks overlap with those identified in the PWAS. CCR, CC chemokine receptor.

MR between social isolation, loneliness and proteins

To infer causality between social isolation, loneliness and the identified proteins and protein modules, we implemented a bidirectional two-sample MR. Genome-wide association study (GWAS) summary statistics were sourced from non-overlapping samples from the UK Biobank. In the forward direction, no protein or protein network was found to be associated with social isolation or loneliness at an FDR

significance threshold, using either *cis*-protein quantitative trait loci (pQTLs) alone or a combination of *cis*- and *trans*-pQTLs (Supplementary Table 10). However, we uncovered significant associations between loneliness and five proteins in the backwards direction using the inverse-variance weighting (IVW) method ($\text{FDR } q < 0.025$) using both *cis*- and *trans*-pQTLs (Fig. 3 and Supplementary Table 11). The results of the sensitivity analyses were consistent with IVW estimates in direction

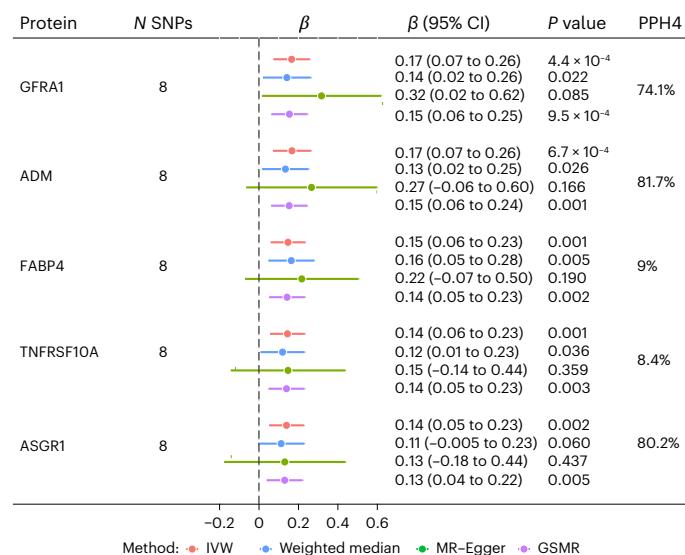


Fig. 3 | Summary of MR and colocalization analysis on the associations between loneliness and proteins and protein modules.

The IVW method was used as the primary analytic approach, complemented by weighted median, MR-Egger and GSMR. The GWAS sample sizes were 297,396 for social isolation, 288,696 for loneliness, 35,327 for GFRA1, 35,385 for ADM, 35,544 for FABP4, 34,842 for TNFRSF10A and 34,842 for ASGR1. All statistical tests were two sided, and FDR correction was applied separately for each direction, as well as for social isolation and loneliness, with significance set at $q < 0.025$. Significant proteins identified by the IVW method are shown. The dots represent beta values and error bars indicate 95% CIs. Support for colocalization was considered strong for PPH4 ≥ 0.8 and medium for $0.5 < \text{PPH4} < 0.8$.

and magnitude. No evidence of heterogeneity (Q statistic, all $P > 0.3$) and horizontal pleiotropy (MR-Egger intercept, all $P > 0.3$; heterogeneity in independent instrument-outlier test, $P > 0.01$) among instrument variables (IVs) was detected. A leave-one-out analyses demonstrated no potentially influential single-nucleotide polymorphisms (SNPs) driving the causal link (Supplementary Fig. 19). However, the causal analysis using summary effect (CAUSE) method could not distinguish a model of causality from correlated pleiotropy for the five proteins that showed significant causal links with loneliness using IVW (Supplementary Table 12).

For the MR-identified proteins in relation to loneliness, we implemented colocalization analysis to ensure that the results were not confounded by linkage disequilibrium. Two proteins, ADM and asialoglycoprotein receptor 1 (ASGR1), exhibited strong evidence of colocalization (posterior probability of hypothesis 4 (PPH4) > 0.8 for one common causal variant) for at least one of the instruments, while GDNF family receptor alpha 1 (GFRA1) showed medium support for colocalization between loneliness and pQTL signals (PPH4 of 0.741) (Fig. 3).

Broad phenotypic associations of MR-identified proteins

We then extended our analysis to explore the broad associations of the MR-identified proteins with diverse blood biomarkers and brain volume. Following Bonferroni correction ($P < 0.05/(229 \times 5) = 4.4 \times 10^{-5}$), we observed extensive associations of these proteins with markers in biochemistry, haematology and metabolomics (Fig. 4a and Supplementary Table 13). Notably, cystatin C showed the strongest association with all of these five proteins. Additionally, all the MR-identified proteins and protein modules demonstrated robust connections with C-reactive protein (CRP) ($t = (13.2, 32.8)$, $N = (40,491, 41,145)$, $P < 9.6 \times 10^{-40}$), a substantial biomarker of systemic inflammation³¹. In brain level, four proteins were significantly associated with volumes in brain regions predominantly located in insula, caudate and frontal cortex, which are known to be involved in interoception, emotional and social

processes^{32–34} (Fig. 4b and Supplementary Table 14). A higher level of ADM was related to lower grey matter volumes in bilateral insula and left caudate (Fig. 4c).

Associations of related proteins with morbidity and mortality

We examined the relationship of the 179 unique proteins and four protein modules linked to social isolation or loneliness with five diseases and mortality using Cox proportional hazards models. With adjustment of demographic, socioeconomical and lifestyle confounders and the first 20 genetic PCs, our findings revealed that 90.2% of these proteins were associated with mortality and over 50% were linked to T2D, CVD and stroke, whereas only 6.6% were associated with dementia ($P < 0.05/(183 \times 6) = 4.6 \times 10^{-5}$; Fig. 5a and Supplementary Table 15). Remarkably, GDF15 emerged as the top protein associated with an increased risk of all diseases and mortality, excluding T2D, while M8 demonstrated the strongest association with T2D (hazard ratio (HR) of 2.14, 95% CI 2.02 to 2.27, $P = 1.6 \times 10^{-146}$, $N = 36,534$). In the case of the MR-identified proteins, we observed significant associations with CVD, T2D, stroke and mortality. ADM was the only one showing an association with dementia (Fig. 5b).

Mediating role of proteins in loneliness and health outcomes

Finally, we delved into the potential mediating role of proteins, which have been implicated as causally linked to loneliness, in the relationship between loneliness and health outcomes. Initially, we estimated the percentage of excess risk mediated (PERM)^{35,36} by each MR-identified protein for the association between loneliness and diseases and mortality through two Cox proportional models. After controlling for demographic, socioeconomical and lifestyle confounders, loneliness was not associated with the incidence of T2D and was therefore excluded from further mediation analyses. The largest effects were observed for mortality (PERM 8.6–16.3%) and CVD (PERM 5.6–8.3%) (Supplementary Table 16). Notably, ADM emerged as the primary mediator linking loneliness to various health outcomes, including CVD (8.3%), dementia (4.4%), stroke (7.8%) and mortality (16.3%). Although this approach provided a quantifiable contribution of proteins and is widely used in epidemiological research^{37,38}, the statistical significance and causal interpretation were unclear. Consequently, we used a counterfactual-based mediation analysis³⁹ to assess the direct and indirect effects in the relationship between loneliness and health outcomes through the MR-identified proteins. All five proteins significantly mediated the association between loneliness and CVD, stroke and mortality after Bonferroni correction (Fig. 5c and Supplementary Table 17). The estimated proportion of the indirect effect to the total effect was comparable to the PERM results. Interestingly, ADM was the only protein that significantly mediated the relationship between loneliness and all four diseases (CVD: indirect effect 9.0×10^{-5} , 95% CI 7×10^{-5} to 1.1×10^{-4} , $P < 1 \times 10^{-16}$, $N = 37,846$; dementia: indirect effect 2.7×10^{-5} , 95% CI 1.6×10^{-5} to 3.7×10^{-5} , $P = 1.4 \times 10^{-6}$, $N = 41,309$; depression: indirect effect 1.8×10^{-5} , 95% CI 7.1×10^{-6} to 2.9×10^{-5} , $P = 0.001$, $N = 37,330$; and stroke: indirect effect 3.5×10^{-5} , 95% CI 2.5×10^{-5} to 4.6×10^{-5} , $P = 5.4 \times 10^{-11}$, $N = 40,555$) and mortality (indirect effect 2.6×10^{-4} , 95% CI 2.3×10^{-4} to 2.9×10^{-4} , $P < 1 \times 10^{-16}$, $N = 41,389$).

Discussion

Understanding the proteomic signatures associated with social isolation and loneliness can provide insight into relevant biological processes, with implications for public health. Leveraging data from 2,920 plasma protein analytes across over 40,000 UK Biobank participants, the present study comprehensively characterized proteins and protein networks related to social isolation and loneliness. The identified proteins were highly interactive and were largely enriched in inflammation, antiviral responses and complement systems. Notably, more than half of these proteins were prospectively linked to CVD,

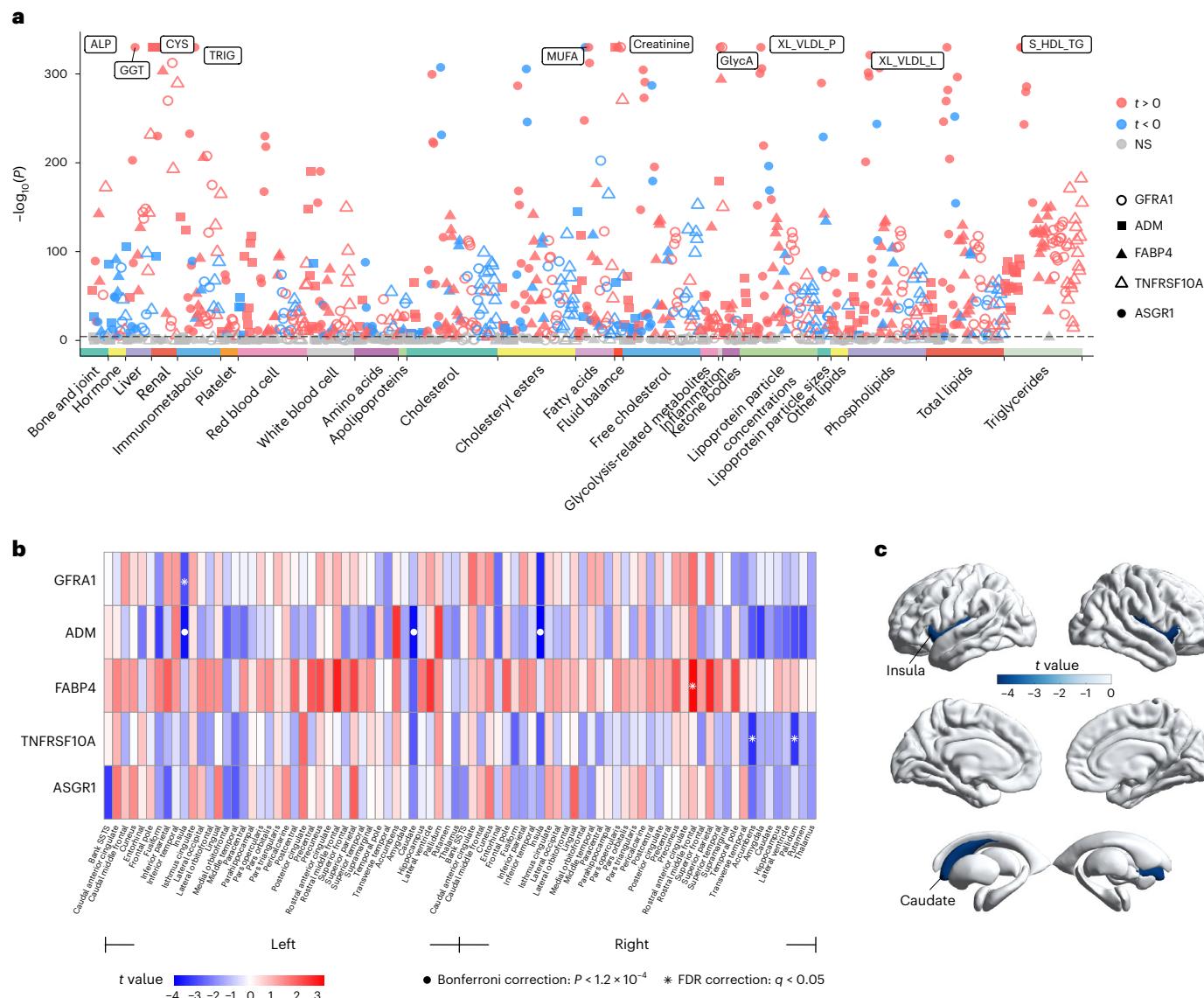


Fig. 4 | Association of MR-identified proteins with blood biomarkers and brain volumes. **a**, The association of MR-identified proteins with biochemical, haematological and metabolic biomarkers. The average sample size for the 229 blood markers analysed was 28,501, ranging from 3,696 to 42,145 (detailed in Supplementary Table 13). Two-sided P values were obtained from linear regression models adjusted for the same covariates as in the PWAS. The x axis represents various blood markers categorized into different groups. The symbols represent different proteins and protein modules, with colours indicating the type of correlation. The dashed line represents the Bonferroni correction ($P < 0.05/(229 \times 5) = 4.4 \times 10^{-5}$). **b**, The association of MR-identified proteins with cortical and subcortical volumes. The average sample size was 5,027, ranging from 4,982 to 5,085 (detailed in Supplementary Table 14). Two-sided P values were derived from linear regression models adjusted for age, sex, imaging collection site, batch, time gap between blood collection

and protein measurement, ethnicity, education level, household income, smoking, alcohol consumption, BMI, the first 20 genetic PCs, intracranial volume and time gap between baseline and imaging collection. The colour bar represents t values, the circles represent significance after Bonferroni correction ($P < 0.05/(84 \times 5) = 1.2 \times 10^{-4}$) and asterisks represent significance after FDR correction ($q < 0.05$) when considering all tested proteins and protein modules simultaneously. Detailed P values are reported in Supplementary Table 14. **c**, Brain volumes significantly associated with the abundance of ADM after FDR correction. The colour bar represents t values. ALP, alkaline phosphatase; GGT, gamma glutamyltransferase; CYS, cystatin C; TRIG, triglycerides; MUFA, monounsaturated fatty acids; GlycA, glycoprotein acetyls; XL_VLDL_P, concentration of very large very low-density lipoproteins (VLDL) particles; XL_VLDL_L, total lipids in very large VLDL; S_HDL_TG, triglycerides in small high-density lipoprotein.

T2D, stroke and mortality during a 14 year follow-up. Using MR, we suggested that loneliness causally contributed to the abundance of five proteins, with two proteins further supported by colocalization. These MR-identified proteins exhibited wide associations with biochemical, haematological and metabolic blood biomarkers, as well as volumes in brain regions involved in interoception and emotional and social processes. Finally, we showed that these MR-identified proteins partly mediated the relationship between loneliness and CVD, stroke and mortality. Collectively, these findings support the idea that social

relationships indirectly influence morbidity and mortality through peripheral physiological pathways.

Social isolation and loneliness, covering separate dimensions of social relationships, can operate independently. The reported behavioural correlations between these constructs range from small to medium ($r = 0.2-0.4$)^{40,41}. They may impact health through distinct pathways, for instance, objective social engagement and loneliness exhibit different associations with neuroimmune markers in older age⁴². Our study revealed a moderate relationship in proteome-wide

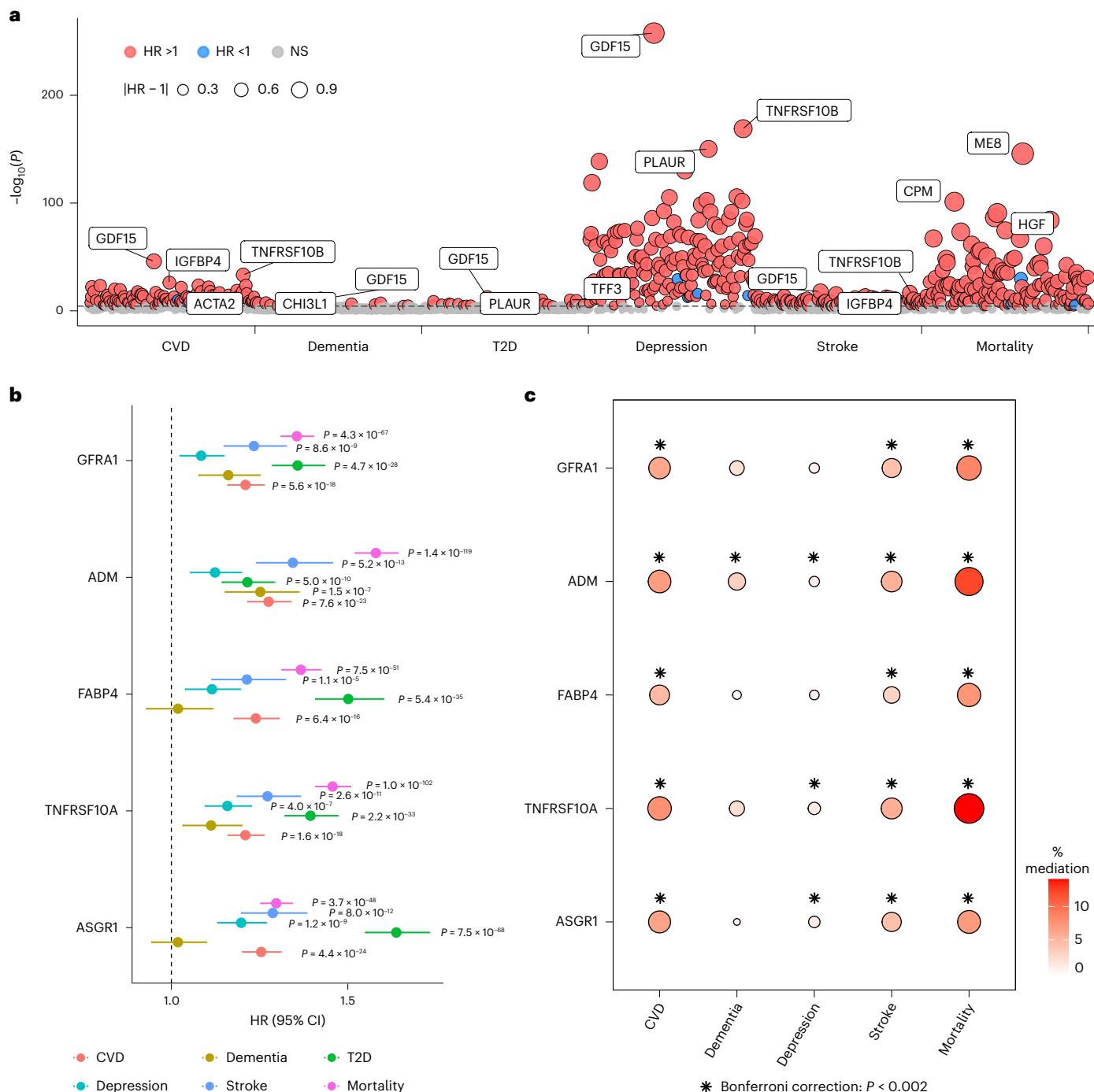


Fig. 5 | Associations between identified proteins and morbidity and mortality.

a, The associations between proteins and protein modules related to social isolation or loneliness with the incidence of five diseases and mortality. The average sample size was 37,808, ranging from 30,232 to 41,396 (detailed in Supplementary Table 15). Two-sided P values were obtained from Cox proportional hazards models adjusted for age, sex, ethnicity, education level, household income, smoking, alcohol consumption, BMI, time gap between blood collection and protein measurement and the first 20 genetic PCs. The top three proteins most strongly associated with each disease and mortality are labelled. The dot colour indicates the type of correlation and the dot size represents the absolute difference between HR and 1. The dashed line represents

the Bonferroni correction ($P < 0.05/(183 \times 6) = 4.6 \times 10^{-5}$). **b**, A forest plot showing the associations between the five MR-identified and health outcomes. The dots represent HRs, with error bars indicating 95% CIs. Significant P values after Bonferroni correction are indicated. **c**, A mediation analysis of MR-identified proteins on the relationship between loneliness and five health outcomes using counterfactual-based analysis. The average sample size was 39,456, ranging from 36,735 to 41,571 (detailed in Supplementary Table 17). The dot size and colour represent the proportion of mediation estimated by indirect effect/(indirect effect + direct effect). The asterisks indicate significance after Bonferroni correction ($P < 0.05/(5 \times 5) = 0.002$). All statistical tests are two sided, and detailed P values are reported in Supplementary Table 17.

associative patterns between social isolation and loneliness ($r = 0.54$). The quantity of proteins significantly linked to social isolation is six times higher than those associated with loneliness. Sensitivity analysis showed that additionally controlling for the other construct had a

minimal impact on proteomic associations, suggesting independence in the proteomic association with each construct. However, there was substantial overlap, with approximately 85% of the proteins associated with loneliness being shared with social isolation. The identified

proteins associated with social isolation or loneliness formed a large interactive network, with immune-related proteins such as IL6 acting as hub proteins. Co-expression network analysis further confirmed shared and distinct modules associated with social isolation and loneliness.

The coregulation between inflammation and social behaviour has been widely discussed in the literature⁴³. Social isolation and loneliness, acting as social stressors, can trigger stress responses, involving the SNS and HPA axis, leading to altered immune function⁴⁴. Indeed, social isolation and loneliness have been shown to be associated with heightened proinflammatory activity, manifested by elevated levels of inflammation biomarkers such as IL6, fibrinogen and CRP^{45–47}. In addition to inflammation, the SNS and HPA axis may also modulate antiviral responses. Older adults experiencing high levels of loneliness displayed an upregulation of proinflammatory genes along with a downregulation of antiviral genes, known as the conserved transcriptional response to adversity (CTRA)⁴⁸. Mechanistic analyses in a macaque model of loneliness confirmed CTRA upregulation preceding increases in loneliness⁴⁹. Another macaque study discovered that social isolation induced CTRA activation and the antiviral gene expression was impeded by enhancing prosocial engagement⁵⁰. However, it should be noted that the majority of research exploring the relationship between social behaviour, antiviral processes and inflammation so far is based on small sample sizes and results are mixed due to methodological heterogeneity¹⁰. Our study, utilizing a biobank-level dataset and various analytical strategies, substantiated the association between social isolation, loneliness and proteins implicated in inflammation and antiviral processes.

Furthermore, both PWAS and network analysis revealed an association between social isolation and proteins linked to complement activation. Although no direct association was found between loneliness and complement in PWAS, loneliness was significantly linked to the protein M3 after FDR correction, which exhibited enrichment in the complement system. Complement acts as a rapid and efficient immune surveillance system, playing a pivotal role in maintaining homeostasis⁵¹, a balance that could be disrupted by social adversity. Recent studies propose that protein components of the complement and coagulation systems could be key pathways in the pathology of psychosis as these pathways can become active years before the onset of psychosis, demonstrating sensitivity to the progression and manifestation of psychotic symptoms⁵². Moreover, we observed that social isolation was uniquely associated with MAPK/ERK signalling. Animal studies have shown that ERK2 conditional knockout mice displayed lower levels of social interactions⁵³. Pair-housing could reverse the detrimental effects of post-stroke social isolation mediated by brain-derived neurotrophic factor via downstream MAPK/ERK signalling⁵⁴. Both human and animal studies suggest that ERK is crucial for pathogenesis, symptomatology and treatment of depression⁵⁵.

Beyond the associative analyses, we integrated proteomic and genetic data to infer the directionality of the relationship between plasma proteins, social isolation and loneliness. Interestingly, we identified significant causal impacts only from loneliness to five proteins, and the robustness of these findings was confirmed through sensitivity analyses, except the CAUSE model. Prior research has established that protein–phenotype associations with both MR and colocalization evidence can predict a higher likelihood of success for a particular target–indication pair, making these proteins more likely to inform clinical applications⁵⁶. In our study, the associations of loneliness with ADM and ASGR1 were strongly supported by colocalization. By examining correlations with other blood biomarkers in the UK Biobank, we discovered that both ADM and ASGR1 were strongly associated with a range of biochemical, haematological and metabolic biomarkers, including CRP, cholesterol and triglycerides. These findings are consistent with previous studies reporting the role of ADM in organizing neuroendocrine responses to stress, including simulating the SNS

and HPA axis⁵⁷, as well as regulating stress hormones⁵⁸. Brain-derived ADM has the potential to control oxytocin release⁵⁹, which is a crucial social peptide⁶⁰. Additionally, variant ASGR1 was associated with low cholesterol and a reduced risk of CVD^{61,62}. We also explored the relationship between the MR-identified proteins and brain volumes, aligning with the growing research focus of characterizing neural mechanisms underlying social relationships⁶³, as related neural mechanisms may translate social experience into health-relevant physiological responses⁴⁴. The strongest association was found between ADM and the insula, a hub for interoceptive mechanisms⁶⁴. Interoception is proposed to function as a vital pathway for the brain–body interactions underlying the loneliness–health link⁶⁵. Other significant associations were observed between plasma ADM levels and the left caudate, a region involved in emotional, reward and social processes^{44,66}.

While the link between social relationships and health is well established, uncovering the underlying biological processes remains a challenge⁹. In this study, we directly investigated the prospective associations between identified proteins, protein modules and health outcomes using longitudinal data from the UK Biobank. All proteins with a significant causal relationship with loneliness were correlated with the incidence of CVD, T2D, stroke and mortality. For example, a one-unit increase in ADM level was associated with a 58% heightened risk of mortality during the follow-up. Furthermore, on average, 7.5% of the association between loneliness and the risk of four diseases and mortality could be attributed to the levels of ADM in plasma based on the Cox-based approach. The mediating effects of ADM on the relationship between loneliness and all five health outcomes were confirmed through mediation analyses based on counterfactuals. Additionally, other MR-identified proteins also showed significant mediating effects for CVD, stroke and mortality. This is the first study to explore the link between loneliness and health outcomes through plasma proteins in humans within a survival context.

However, it is important to note that the lack of an association between plasma proteins, social isolation and loneliness in MR analyses does not necessarily preclude a mechanistic role for candidate proteins in the connection between social relationships and health. In fact, over half of the proteins and protein modules exhibiting a significant phenotypic correlation with social isolation or loneliness demonstrated a noteworthy association with CVD, T2D, stroke and mortality. Remarkably, GDF15 emerged as the top-risk protein for CVD, stroke, dementia, depression and mortality. A recent study identified GDF15 plasma levels as the top robust predictor across 14 categories of disorders and all-cause mortality⁶⁷. However, GDF15 might also exert anti-inflammatory effects by inhibiting macrophage activation, thus potentially playing a protective role⁶⁸. The disparate roles of GDF15 might be mediated by different mechanisms that are not fully understood. Additionally, consistent with our results, GDF15 was not found to be causally associated with diseases such as hypertension, diabetes, heart disease, stroke and cancer⁶⁹. Therefore, the increases in GDF15 observed in pathologies might be a consequence of these diseases, potentially resulting from inflammation caused by these diseases, rather than a cause⁷⁰. It is also possible that associations with disease development are potentially compensatory⁷¹. These pieces of evidence may imply that GDF15 is not specific to a particular disease, yet this does not rule out its potential mechanistic relevance to neuroinflammation or other processes pertinent to the link between social relationships and health.

The following considerations should be kept in mind when interpreting our findings. First, the measurements of loneliness and social isolation within the UK Biobank do not differentiate between acute and chronic conditions. While acute social isolation/loneliness, activating the SNS and HPA axis, may be adaptive in the short term, persistent activation of these systems is associated with increased inflammatory activity and implicated in the progression of multiple disease states and psychopathologies⁷². Future research should aim to distinguish

between these conditions. Second, we observed a notable decrease in the associations between plasma proteins and measures of social relationships, particularly with loneliness, when accounting for baseline depressive symptoms. Previous research indicated that 75% of the connection between loneliness and dementia³⁵, as well as 66% of the link between loneliness and mortality⁷³, could be attributed to depressive symptoms. Additionally, there is a high genetic correlation between loneliness and depression ($r = 0.63$), nearly equivalent the genetic correlation between loneliness and social isolation ($r = 0.65$)⁷⁴, indicating a significant co-occurrence of these phenotypes. Third, although the Olink proteomics platform utilized in this study offers a comprehensive measurement of the human proteome, not all circulating proteins were measured. Additionally, we have characterized plasma proteomic signatures of social isolation and loneliness, but we do not know the relationship with protein levels in other tissues as protein levels are known to differ between cell types⁷⁵. Fourth, consistent causal effects of loneliness on the five proteins were inferred using different methods except for CAUSE, indicating the possibility of correlated pleiotropy. Correlated pleiotropy is common and may frequently contribute to false positives in MR studies⁷⁶. Further investigations in larger, more powerful datasets will be needed to determine whether the observed associations are due to causality rather than pleiotropy. Last, an important limitation of this study is the lack of external validation, given that social isolation and loneliness are not typically measured variables. Nevertheless, cross-validation in the UK Biobank and sensitivity analyses supported the robustness of our findings. Future validation in an independent dataset is essential.

In conclusion, this is the first study delineating robust and comprehensive plasma proteomic signatures associated with social isolation and loneliness. The plasma proteome can help bridge the link between social relationships and morbidity and mortality. Comprehending the biology underlying the impact of social relationships on health, particularly the peripheral changes preceding disease, may provide new opportunities for targeted prevention and for effective intervention.

Methods

Study cohort

The UK Biobank is a population-based cohort that involves over 500,000 individuals aged 40–69 years old recruited from 22 centres across the United Kingdom between 2006 and 2010⁷⁷. Participant data include genome-wide genotyping, magnetic resonance imaging, electronic health record linkage, blood and urine biomarkers, as well as various other phenotypic endpoints. All participants in the UK Biobank provided informed consent, and ethical approval was obtained from the National Information Governance Board for Health and Social Care and the North West Multi-Centre Research Ethics Committee (ref. 11/NW/0382).

Defining social isolation and loneliness

Social isolation and loneliness were calculated from scales that have been widely used in previous studies^{35,73}. Social isolation was assessed through three dichotomous questions (living alone (1 = yes), social contact (1 = less than monthly) and participation in social activities (1 = less than weekly)), with an individual defined as socially isolated if scoring 2 or 3. Loneliness was evaluated using two binary items akin to questions in the revised University of California, Los Angeles loneliness scale⁷⁸: often feeling lonely (1 = yes) and frequency of confiding in close people (1 = less than once every few months). An individual was classified as lonely if scoring 2.

Proteomic measurements

The UKB-PPP consortium conducted proteomic profiling on blood plasma samples collected from 53,026 participants at baseline, using the Olink Explore 3072 Proximity Extension Assay between April 2021 and February 2022. The subcohort comprised a randomly selected

subset of 45,507 participants at baseline, 6,229 individuals preselected by the UKB-PPP consortium and 1,290 individuals participating in the COVID-19 repeat-imaging study at multiple visits. Consortium-selected and COVID-19 imaging participants showed widespread differences such as demographics and biochemistry markers, reflecting their non-random sampling, whereas the randomly selected baseline participants remained highly representative of the overall UK Biobank⁷⁹. Following stringent quality control (QC) procedures (refer to biobank.ndph.ox.ac.uk/ukb/ukb/docs/PPP_Phase_1_QC_dataset_companion_doc.pdf for detailed information), 2,923 unique proteins were measured across eight panels containing cardiometabolic (I and II), inflammation (I and II), neurology (I and II) and oncology (I and II). The inter- and intraplate coefficients of variation for all Olink panels were below 20% and 10%, respectively. Due to the QC process, different participants had missing values for different proteins, resulting in varying sample sizes for individual proteins. After excluding proteins with missing data exceeding 50% (GLIPR1, NPM1 and PCOLCE), a total of 2,920 unique proteins were included. We utilized the reported normalized protein expression values, and normalized protein expression underwent inverse-rank normalization before analysis⁷⁹.

Genotyping and QC

Genotype imputation data were available for 487,409 participants (version 3, released in May 2017). Two closely related arrays, the Applied Biosystems UK BiLEVE Axiom by Affymetrix (now part of Thermo Fisher Scientific) and the Applied Biosystems UK Biobank Axiom Array, were employed to genotype over 800,000 markers with good genome-wide coverage⁸⁰. We removed samples that had mismatch between genetically inferred and self-reported sex, high genotype missingness or extreme heterozygosity, and those with more than ten putative third-degree relatives. Our analysis was restricted to subjects identified as Caucasians through principal component analysis. We further filtered out SNPs with call rates <95%, minor allele frequency <0.5% and deviation from the Hardy–Weinberg equilibrium with $P < 1 \times 10^{-6}$, using PLINK 2 (ref. 81). After QC procedures, we retained a total of 9,910,057 SNPs and 337,138 participants.

Structural magnetic resonance imaging data

All neuroimaging data were acquired, preprocessed, quality controlled and made available by the UK Biobank (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf). Neuroimaging data were collected with a standard Siemens Skyra 3T scanner with a 32-channel head coil. T1-weighted images were processed with FreeSurfer v6 to derive measurements of cortical and subcortical volumes. The Qoala-T approach was used to check FreeSurfer outputs, with additional manual checks performed for outputs close to the threshold. Any outputs that did not pass QC were excluded^{82,83}. A total of 68 cortical regions parcellated by the Desikan–Killiany atlas⁸⁴ and 16 subcortical regions defined by the aseg atlas⁸⁵ were used in the current study.

Blood biomarkers

The UK Biobank blood sample collection was undertaken at baseline. Blood biochemistry data (category 17518) were derived from approximately 480,000 participants. Detailed QC procedures can be found at https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/biomarker_issues.pdf. We categorized 30 biochemical markers into five groups, covering liver, renal, hormone, bone and joint, and immunometabolism. Meanwhile, blood count data (category 100081), encompassing 31 haematological markers related to red blood cell, white blood cell and platelet, were collected from the same number of participants. In-depth information about the haematology analysis is available at <https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/haematology.pdf>. Metabolic biomarkers were measured in randomly selected EDTA plasma samples from approximately 280,000 UK Biobank participants, using a high-throughput NMR-based metabolic biomarker profiling

platform developed by Nightingale Health Ltd. Further details on the processing and QC are described previously⁸⁶. In this study, we utilized absolute concentrations of 168 biomarkers, which were categorized into 16 groups such as amino acids, apolipoproteins and cholesterol. The R package ‘ukbnmr’⁸⁷ was employed to remove technical variation in NMR data. All 229 blood biomarkers were inverse-rank normalized before analyses.

Ascertainment of health outcomes

The identification of five diseases (CVD: International Classification of Diseases tenth codes of I20, I21, I22, I23, I24 and I25; all-cause dementia: F00, F01, F02, F03 and G30; T2D: E11; depression: F32 and F33; and stroke: I60, I61, I63 and I64) relied on the ‘first occurrence’ data fields generated by the UK Biobank (<https://biobank.ndph.ox.ac.uk/showcase/label.cgi?id=1712>). These data were ascertained through a combination of primary care, hospital in-patient, death register and self-reported information. The date of diagnosis was set as the earliest date of diagnosis, regardless of the source used. Prevalent cases were defined as diagnoses within the first 3 years of follow-up or self-reported cases at baseline, and these cases were excluded to mitigate potential reverse-causation bias. Follow-up for deaths started at inclusion in the UK Biobank study (from national death registers) and ended on 30 November 2022 or upon death, for all participants.

Covariates

Covariates were selected based on literature and data availability. Demographic factors included age, sex, assessment centre, ethnicity (white, mixed, Asian, Black and other), BMI, household income (<£31,000 or ≥£31000) and education level. Education level was grouped into four categories reflecting similar years of education (college/university degree, education to age 18 years or above, education to age 16 years qualifications and no qualifications)⁸⁸. Current smoking status and alcohol intake frequency (<3 or ≥3 times per week) were obtained by self-rated questions. Depressive symptoms over the past 2 weeks were assessed using four items from the Patient Health Questionnaire⁸⁹ (depressed mood, uninterest/absence of enthusiasm, tenseness/restlessness and tiredness/lethargy). Physical activity was assessed using the International Physical Activity Questionnaire short form⁹⁰, which includes six questions on the frequency and duration of walking, moderate-intensity exercise and vigorous exercise. Following the scoring protocol⁹¹, responses of ‘unable to walk’ were recoded to 0, while ‘do not know’ and ‘prefer not to answer’ were treated as missing data. Bouts of activity lasting less than 10 min duration were not counted and activity bouts of longer than 4 h were truncated. Weekly minutes of each activity category were multiplied by the corresponding metabolic equivalent (MET) values (that is, 3.3 for walking, 4.0 for moderate physical activity and 8.0 for vigorous physical activity)⁹², then summed to obtain the total MET minutes representing the energy expended during physical activity. Technical factors for plasma proteins included batch and time gap between blood collection and protein measurement. The first 20 genetic PCs were controlled for to mitigate potential population stratification.

Statistical analysis

PWAS. Logistic regression was performed with social isolation or loneliness as the outcome and protein level as the predictor. Initially, a simple model was employed, controlling for age, sex, site, technical factors and the first 20 genetic PCs. Subsequently, a fully adjusted model was examined, additionally controlling for ethnicity, education level, household income, smoking, alcohol consumption and BMI. The fully adjusted model was used for all primary analyses.

Several sensitivity analyses were conducted to assess the robustness of these findings: (1) To address the limitations of dichotomous variables potentially losing statistical power, raw scores were used and modelled with ordered logistic regression, controlling for all covariates

in the primary model. (2) To address potential bias stemming from the UKB-PPP subcohort composition, the primary analyses were replicated exclusively within the randomly selected subset. (3) PWAS was conducted specifically in Caucasians to further mitigate the possibility of population stratification. (4) Sex and age differences were evaluated by incorporating interaction terms between sex or age and protein levels into logistic models, adjusting for the same covariates as in the primary model. Additionally, subgroup PWAS analyses were conducted separately for males and females, as well as for younger (<60 years) and older (≥60 years) individuals, to provide further insights. (5) Depressive symptoms and physical activity were included to assess their respective impacts on the relationship between social isolation/loneliness and proteins. Social isolation and loneliness were simultaneously included in the model to test for mutual influence. The reduced percentage of OR was calculated to quantify the influence of additional factors. (6) Participants were categorized into four groups: neither isolated nor lonely (SI-LO-), socially isolated but not lonely (SI+LO-), not isolated but lonely (SI-LO+) and socially isolated and lonely (SI+LO+). Multinomial logistic regression was performed with all covariates from the primary model, using the SI-LO- group as the reference. A likelihood ratio test was used to estimate the overall association of protein levels compared with the model with covariates only. Significant proteins related to a specific group required both a significant model fit and a significant coefficient after Bonferroni correction ($P < 0.05/2920 = 1.7 \times 10^{-5}$). (7) Cross-validation was performed by randomly dividing socially isolated, lonely and control participants to ensure similar sample sizes for each group in the two split datasets. Fully adjusted PWAS analyses were performed separately on the two datasets, and this process was repeated 100 times to ensure reliability.

Protein co-expression network analysis. Participants with missing protein data exceeding 50% were excluded, and the remaining missing proteins were imputed using the k -nearest imputation function in the R package ‘impute’. Employing the full set of 2,920 proteins from 46,850 participants, we utilized Netboost⁹³, a dimension-reduction procedure that extends the widely used weighted gene co-expression network analysis method⁹⁴, to identify co-expressed protein networks (modules). The procedure involved three steps: first, we calculated the boosting-based filter and a sparse distance matrix based on Spearman correlation. Then, protein modules were identified through sparse hierarchical clustering and the dynamic tree cut procedure. Finally, we aggregated information in the modules using the first PC. Consistent with prior network-based proteomic analyses⁹⁵, we set the minimum module size to 20, applied a Spearman filter method with a soft power of $\beta = 2$ and an ‘unsigned’ network approach in clustering. Additionally, we compared protein networks constructed with different soft powers (that is, 3 and null) against these parameters. The association between the expression of each protein module and social isolation and loneliness was examined using logistic regression, adjusting for covariates in the fully adjusted model. To assess the consistency between the identified protein networks and the proteins associated with social isolation or loneliness identified through PWAS, two-sided Fisher’s exact tests were conducted to determine whether the identified proteins were significantly enriched in the identified networks using all proteins in the module, as well as the top 20% and top 10% of proteins that exhibit the most significant correlation with the corresponding module eigengene.

PPI network and functional enrichment analysis. The PPI network of the proteins of interest was constructed using STRING²⁸ with default settings and visualized by Cytoscape⁹⁶. The cytoHubba plugin²⁹ in Cytoscape was utilized to identify hub proteins by the MCC method.

To assess the biological relevance of the identified protein sets, gprofiler2 (ref. 97) was employed to calculate functional enrichments for Gene Ontology (GO)⁹⁸ biological processes and molecular functions, Kyoto Encyclopedia of Genes and Genomes (KEGG)⁹⁹, Reactome¹⁰⁰ and

WikiPathways¹⁰¹. Proteins associated with social isolation/loneliness after FDR correction ($q < 0.05$, corrected for two constructs simultaneously) were used as the test set, with all proteins measured in the UK Biobank ($N = 2,920$) used as the background set. The significance of enrichment was indicated by the P value of the hypergeometric test, followed by FDR correction. We considered results with FDR $q < 0.05$ as statistically significant.

MR. We applied bidirectional MR to investigate the causal relationship between identified proteins and protein networks with social isolation and loneliness. The GWAS on plasma protein abundance, using protein level for individual proteins and module eigengene for protein networks, was conducted in Caucasian participants with quality-controlled genotyping data from the UK Biobank. To avoid bias due to participant overlap, separate GWASs for social isolation ($N = 297,396$) and loneliness ($N = 288,696$) were performed using a distinct set of Caucasian participants, excluding those involved in the protein-related GWAS. Age, sex and the first 20 genetic PCs were adjusted for in all GWAS analyses, while protein technical factors were additionally accounted for in GWAS analyses of proteins and protein modules. GWAS analyses were performed through PLINK 2 (ref. 81) and GCTA¹⁰².

For the forward direction (protein to social isolation/loneliness), we selected pQTLs with genome-wide significance ($P < 5 \times 10^{-8}$) as IV and applied clumping (1,000 kb distance, maximum linkage disequilibrium (LD) r^2 of 0.001). pQTLs were categorized into *cis*-pQTLs (± 1 MB window of the gene encoding the target protein) and *trans*-pQTLs (outside the ± 1 MB window). Two types of MR analyses were conducted using *cis*-pQTLs alone and both *cis*- and *trans*-pQTLs. For the backward direction (social isolation/loneliness to protein), due to a limited number of SNPs reaching genome-wide significance after clumping, we relaxed the P value threshold to 1×10^{-6} to extract genetic instruments for social isolation and loneliness. We utilized IVW as the primary analytic approach, with the Wald ratio test employed if only one instrument was available. To account for pleiotropy, three sensitivity analyses were conducted: (1) weighted median¹⁰³; (2) MR-Egger¹⁰⁴, which adds an intercept to the IVW regression to exclude confounding from uncorrelated pleiotropy; and (3) generalized summary data-based MR (GSMR) with the heterogeneity in independent instrument-outlier method¹⁰⁵ to identify and remove SNPs with evidence for significant uncorrelated pleiotropy ($P < 0.01$). For GSMR, we used 10,000 randomly selected unrelated samples from the UK Biobank as a reference to determine LD patterns. For significant MR results, CAUSE⁷⁶ was performed with a large set of LD-pruned SNPs ($r^2 < 0.01$ with an arbitrary $P \leq 1 \times 10^{-3}$) to further consider correlated pleiotropy. Additionally, Cochran's Q -test¹⁰⁶ was used to determine IV heterogeneity. A leave-one-out analysis was performed to assess the influence of individual SNPs on MR estimation. FDR correction was applied to the results of each direction and for social isolation and loneliness separately, with FDR $q < 0.025$ considered significant. MR analyses were performed utilizing the R packages 'TwoSampleMR'¹⁰⁷, 'gsmr2'¹⁰⁵ and 'cause'⁷⁶.

Colocalization analysis. The results that survived the multiple testing threshold on MR analysis were evaluated using a colocalization analysis to estimate the posterior probability (PP) of each genomic locus containing a single variant affecting both the protein and the phenotype. The analysis was based on a Bayesian model that assesses the probability of a shared causal variant (PPH4) or distinct causal variants (PPH3)¹⁰⁸. Variants within a ± 500 kb window around the significant SNP ($P < 5 \times 10^{-8}$) with the smallest P value were included. Colocalization was performed using with default priors (PP of initial trait association (p_1) is 1×10^{-4} , PP of second trait association (p_2) is 1×10^{-4} and prior probability of shared causal variant across two traits (p_{12}) is 1×10^{-5}). Two signals were considered to have a strong support of colocalization if the PPH4 ≥ 0.8 . Medium colocalization indication was defined as $0.5 < \text{PPH4} < 0.8$. The analysis was performed using the R package 'coloc'¹⁰⁸.

Associations to other blood biomarkers and brain volumes. Linear regression was used to investigate the associations between proteins and protein modules related to social isolation/loneliness and various blood biochemical, haematological and metabolic markers, as well as cortical and subcortical volumes. The same covariates as in the primary model were controlled for. Additionally, intracranial volume and the time gap between baseline and imaging collection were considered in neuroimaging analyses.

Cox proportional hazards model. Associations between the abundance of social isolation or loneliness associated proteins and five diseases (CVD, dementia, T2D, depression and stroke) and mortality were investigated using Cox proportional hazards models. The proportional hazard assumptions were checked using Schoenfeld residuals and no major violations were observed. All models were adjusted for age, sex, ethnicity, education level, household income, smoking, alcohol consumption, BMI, time gap between blood collection and protein measurement, and the first 20 genetic PCs.

Mediation analysis. To support the reliability of our results, we used two approaches to estimate the contribution of proteins explaining the relationship between social isolation/loneliness and health outcomes. First, we estimated the PERM using two Cox proportional hazards models, with the formula $(\text{HR}_{(\text{age}, \text{sex}, \text{site}, \text{ethnicity}, \text{education level}, \text{household income}, \text{smoking}, \text{alcohol consumption and BMI adjusted})} - \text{HR}_{(\text{age}, \text{sex}, \text{site}, \text{ethnicity}, \text{education level}, \text{household income}, \text{smoking}, \text{alcohol consumption, BMI and protein adjusted})}) / (\text{HR}_{(\text{age}, \text{sex}, \text{site}, \text{ethnicity}, \text{education level}, \text{household income}, \text{smoking}, \text{alcohol consumption and BMI adjusted})} - 1) \times 100$ (refs. 35,36). Given the limitations of the Cox-based mediation analysis in inferring statistical significance and offering a causal interpretation, we subsequently adopted a counterfactual-based mediation analysis¹⁰⁹. In this study, we employed a straightforward procedure based on marginal structural models that directly parameterize the natural direct and indirect effects of interest³⁹. Based on the results of MR analysis, we found a significant causal relationship between loneliness and protein levels. Therefore, loneliness was treated as the exposure and protein level were used as the mediator. In the mediator model, we included all the covariates used in the PWAS (that is, age, sex, site, batch, time gap between blood collection and protein measurement, ethnicity, education level, household income, smoking, alcohol consumption, BMI and the first 20 genetic PCs). In the outcome model, all covariates except protein technical factors and genetic PCs were adjusted for. The Aalen additive hazard model¹¹⁰ was utilized for the survival outcome. CIs were computed as the estimate of the natural direct or indirect effect plus/minus 1.96 times a robust standard error.

Multiple comparison correction. All reported P values in this study are two sided, except for the hypergeometric test used in the enrichment analysis. We provided results following both Bonferroni and FDR corrections. The Bonferroni correction was calculated as 0.05 divided by the total number of tests, encompassing both social isolation and loneliness as predictors. FDR correction was applied to all analyses simultaneously, with significance set at $q < 0.05$. For MR analysis, due to the substantial difference in the number of proteins associated with social isolation and loneliness, separate FDR corrections were applied for each, with significance set at $q < 0.025$.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data used in the present study are available from the UK Biobank (<https://www.ukbiobank.ac.uk>) with restrictions applied. Data were used under licence and are thus not publicly available. Details regarding registration for data access can be found at <http://www.ukbiobank.ac.uk/register-apply/>. The data used in this study were accessed from

the UK Biobank under the application number 19542. GWAS summary statistics used can be found via the figshare website at https://figshare.com/projects/GWAS_summary_data/224229 (ref. 11). European ancestry reference data from the 1000 Genomes Project can be found via GitHub at <https://github.com/getian107/PRScsx?tab=readme-ov-file>.

Code availability

Custom scripts for the analyses have been made available via the following GitHub repository at https://github.com/chunshen617/Proteomics_loneliness.

References

- House, J. S., Landis, K. R. & Umberson, D. Social relationships and health. *Science* **241**, 540–545 (1988).
- Holt-Lunstad, J. A pandemic of social isolation? *World Psychiatry* **20**, 55–56 (2021).
- Holt-Lunstad, J., Smith, T. B. & Layton, J. B. Social relationships and mortality risk: a meta-analytic review. *PLOS Med.* **7**, e1000316 (2010).
- Wang, F. et al. A systematic review and meta-analysis of 90 cohort studies of social isolation, loneliness and mortality. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-023-01617-6> (2023).
- Holt-Lunstad, J. Why social relationships are important for physical health: a systems approach to understanding and modifying risk and protection. *Annu. Rev. Psychol.* **69**, 437–458 (2018).
- Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T. & Stephenson, D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. *Perspect. Psychol. Sci.* **10**, 227–237 (2015).
- Capitanio, J. P., Cacioppo, S. & Cole, S. W. Loneliness in monkeys: neuroimmune mechanisms. *Curr. Opin. Behav. Sci.* **28**, 51–57 (2019).
- Cacioppo, J. T., Cacioppo, S., Capitanio, J. P. & Cole, S. W. The neuroendocrinology of social isolation. *Annu. Rev. Psychol.* **66**, 733–767 (2015).
- Snyder-Mackler, N. et al. Social determinants of health and survival in humans and other animals. *Science* **368**, eaax9553 (2020).
- Smith, K. J., Govey, S., Riddell, N. E., Kontari, P. & Victor, C. The association between loneliness, social isolation and inflammation: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* **112**, 519–541 (2020).
- Leschak, C. J. & Eisenberger, N. I. Two distinct immune pathways linking social relationships with health: inflammatory and antiviral processes. *Psychosom. Med.* **81**, 711–719 (2019).
- Hawley, L. C. & Cacioppo, J. T. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. *Ann. Behav. Med.* <https://doi.org/10.1007/s12160-010-9210-8> (2010).
- Santos, R. et al. A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).
- Salinas, J. et al. Associations between social relationship measures, serum brain-derived neurotrophic factor, and risk of stroke and dementia. *Alzheimers Dement.* **3**, 229–237 (2017).
- Dunbar, R. I. M. & Shultz, S. Evolution in the social brain. *Science* **317**, 1344–1347 (2007).
- Zovetti, N., Rossetti, M. G., Perlini, C., Brambilla, P. & Bellani, M. Neuroimaging studies exploring the neural basis of social isolation. *Epidemiol. Psychiatr. Sci.* **30**, e29 (2021).
- Hodgson, S., Watts, I., Fraser, S., Roderick, P. & Dambha-Miller, H. Loneliness, social isolation, cardiovascular disease and mortality: a synthesis of the literature and conceptual framework. *J. R. Soc. Med.* **113**, 185–192 (2020).
- Valtorta, N. K., Kanaan, M., Gilbody, S., Ronzi, S. & Hanratty, B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. *Heart* **102**, 1009–1016 (2016).
- Kuiper, J. S. et al. Social relationships and risk of dementia: a systematic review and meta-analysis of longitudinal cohort studies. *Ageing Res. Rev.* **22**, 39–57 (2015).
- Livingston, G. et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* **396**, 413–446 (2020).
- Hackett, R. A., Hudson, J. L. & Chilcot, J. Loneliness and type 2 diabetes incidence: findings from the English Longitudinal Study of Ageing. *Diabetologia* **63**, 2329–2338 (2020).
- Lee, S. L. et al. The association between loneliness and depressive symptoms among adults aged 50 years and older: a 12-year population-based cohort study. *Lancet Psychiatry* **8**, 48–57 (2021).
- Smith, R. W. et al. Social isolation and risk of heart disease and stroke: analysis of two large UK prospective studies. *Lancet Public Health* **6**, e232–e239 (2021).
- Rochette, L., Zeller, M., Cottin, Y. & Vergely, C. Insights into mechanisms of GDF15 and receptor GFRAL: therapeutic targets. *Trends Endocrinol. Metab.* **31**, 939–951 (2020).
- Abifadel, M. et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat. Genet.* **34**, 154–156 (2003).
- Lu, J., Chatterjee, M., Schmid, H., Beck, S. & Gawaz, M. CXCL14 as an emerging immune and inflammatory modulator. *J. Inflamm.* **13**, 1 (2016).
- Desquibet, L. & Mariotti, F. Dose-response analyses using restricted cubic spline functions in public health research. *Stat. Med.* **29**, 1037–1057 (2010).
- Szklarczyk, D. et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **51**, D638–D646 (2022).
- Chin, C.-H. et al. cytoHubba: identifying hub objects and subnetworks from complex interactome. *BMC Syst. Biol.* **8**, S11 (2014).
- Emilsson, V. et al. Co-regulatory networks of human serum proteins link genetics to disease. *Science* **361**, 769–773 (2018).
- Saito, K. & Kihara, K. C-reactive protein as a biomarker for urological cancers. *Nat. Rev. Urol.* **8**, 659–666 (2011).
- Hare, T. A., Tottenham, N., Davidson, M. C., Glover, G. H. & Casey, B. J. Contributions of amygdala and striatal activity in emotion regulation. *Biol. Psychiatry* **57**, 624–632 (2005).
- Frith, C. D. & Frith, U. Social cognition in humans. *Curr. Biol.* **17**, R724–R732 (2007).
- Singer, T., Critchley, H. D. & Preuschoff, K. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* **13**, 334–340 (2009).
- Shen, C. et al. Associations of social isolation and loneliness with later dementia. *Neurology* **99**, e164–e175 (2022).
- Lin, D. Y., Fleming, T. R. & De Gruttola, V. Estimating the proportion of treatment effect explained by a surrogate marker. *Stat. Med.* **16**, 1515–1527 (1997).
- Lu, Y. et al. Metabolic mediators of the effects of body-mass index, overweight, and obesity on coronary heart disease and stroke: a pooled analysis of 97 prospective cohorts with 1.8 million participants. *Lancet* **383**, 970–983 (2014).
- Etemadi, A. et al. Mortality from different causes associated with meat, heme iron, nitrates, and nitrites in the NIH-AARP Diet and Health Study: population based cohort study. *Br. Med. J.* **357**, j1957 (2017).
- Lange, T., Vansteelandt, S. & Bekaert, M. A simple unified approach for estimating natural direct and indirect effects. *Am. J. Epidemiol.* **176**, 190–195 (2012).
- Coyle, C. E. & Dugan, E. Social isolation, loneliness and health among older adults. *J. Aging Health* **24**, 1346–1363 (2012).
- Perissinotto, C. M. & Covinsky, K. E. Living alone, socially isolated or lonely—what are we measuring? *J. Gen. Intern. Med.* **29**, 1429–1431 (2014).
- Walker, E., Ploubidis, G. & Fancourt, D. Social engagement and loneliness are differentially associated with neuro-immune markers in older age: time-varying associations from the English Longitudinal Study of Ageing. *Brain. Behav. Immun.* **82**, 224–229 (2019).

43. Eisenberger, N. I., Moieni, M., Inagaki, T. K., Muscatell, K. A. & Irwin, M. R. In sickness and in health: the co-regulation of inflammation and social behavior. *Neuropsychopharmacology* **42**, 242–253 (2017).
44. Eisenberger, N. I. & Cole, S. W. Social neuroscience and health: neurophysiological mechanisms linking social ties with physical health. *Nat. Neurosci.* **15**, 669–674 (2012).
45. Ford, E. S., Loucks, E. B. & Berkman, L. F. Social integration and concentrations of C-reactive protein among US adults. *Ann. Epidemiol.* **16**, 78–84 (2006).
46. Heffner, K. L., Waring, M. E., Roberts, M. B., Eaton, C. B. & Gramling, R. Social isolation, C-reactive protein, and coronary heart disease mortality among community-dwelling adults. *Soc. Sci. Med.* **72**, 1482–1488 (2011).
47. Nersesian, P. V. et al. Loneliness in middle age and biomarkers of systemic inflammation: findings from midlife in the United States. *Soc. Sci. Med.* **209**, 174–181 (2018).
48. Cole, S. W. et al. Social regulation of gene expression in human leukocytes. *Genome Biol.* **8**, R189 (2007).
49. Cole, S. W. et al. Myeloid differentiation architecture of leukocyte transcriptome dynamics in perceived social isolation. *Proc. Natl Acad. Sci. USA* **112**, 15142–15147 (2015).
50. Cole, S. W. et al. The Type I interferon antiviral gene program is impaired by lockdown and preserved by caregiving. *Proc. Natl Acad. Sci. USA* **118**, e2105803118 (2021).
51. Ricklin, D., Hajishengallis, G., Yang, K. & Lambris, J. D. Complement: a key system for immune surveillance and homeostasis. *Nat. Immunol.* **11**, 785–797 (2010).
52. Heurich, M., Föcking, M., Mongan, D., Cagney, G. & Cotter, D. R. Dysregulation of complement and coagulation pathways: emerging mechanisms in the development of psychosis. *Mol. Psychiatry* **27**, 127–140 (2022).
53. Satoh, Y. et al. ERK2 contributes to the control of social behaviors in mice. *J. Neurosci.* **31**, 11953–11967 (2011).
54. Verma, R. et al. Reversal of the detrimental effects of post-stroke social isolation by pair-housing is mediated by activation of BDNF-MAPK/ERK in aged mice. *Sci. Rep.* **6**, 25176 (2016).
55. Wang, J. Q. & Mao, L. The ERK pathway: molecular mechanisms and treatment of depression. *Mol. Neurobiol.* **56**, 6197–6205 (2019).
56. Zheng, J. et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat. Genet.* **52**, 1122–1131 (2020).
57. Shan, J. & Krukoff, T. L. Intracerebroventricular adrenomedullin stimulates the hypothalamic–pituitary–adrenal axis, the sympathetic nervous system and production of hypothalamic nitric oxide. *J. Neuroendocrinol.* **13**, 975–984 (2001).
58. Taylor, M. M. & Samson, W. K. A possible mechanism for the action of adrenomedullin in brain to stimulate stress hormone secretion. *Endocrinology* **145**, 4890–4896 (2004).
59. White, M. M. & Samson, W. K. A possible relationship between brain-derived adrenomedullin and oxytocin in the regulation of sodium balance. *J. Endocrinol.* **203**, 253–262 (2009).
60. Dunbar, R. I. M. The anatomy of friendship. *Trends Cogn. Sci.* **22**, 32–51 (2018).
61. Nioi, P. et al. Variant ASGR1 associated with a reduced risk of coronary artery disease. *N. Engl. J. Med.* **374**, 2131–2141 (2016).
62. Wang, J.-Q. et al. Inhibition of ASGR1 decreases lipid levels by promoting cholesterol excretion. *Nature* **608**, 413–420 (2022).
63. Delgado, M. R., Fareri, D. S. & Chang, L. J. Characterizing the mechanisms of social connection. *Neuron* **111**, 3911–3925 (2023).
64. Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A. & Dolan, R. J. Neural systems supporting interoceptive awareness. *Nat. Neurosci.* **7**, 189–195 (2004).
65. Quadt, L., Esposito, G., Critchley, H. D. & Garfinkel, S. N. Brain–body interactions underlying the association of loneliness with mental and physical health. *Neurosci. Biobehav. Rev.* **116**, 283–300 (2020).
66. Xiong, Y., Hong, H., Liu, C. & Zhang, Y. Q. Social isolation and the brain: effects and mechanisms. *Mol. Psychiatry* <https://doi.org/10.1038/s41380-022-01835-w> (2022).
67. You, J. et al. Plasma proteomic profiles predict individual future health risk. *Nat. Commun.* **14**, 7817 (2023).
68. Pence, B. D. Growth differentiation factor-15 in immunity and aging. *Front. Aging* **3**, 837575 (2022).
69. Tanaka, T. et al. Plasma proteomic biomarker signature of age predicts health and life span. *eLife* **9**, e61073 (2020).
70. Luan, H. H. et al. GDF15 is an inflammation-induced central mediator of tissue tolerance. *Cell* **178**, 1231–1244.e11 (2019).
71. Wang, D. et al. GDF15: emerging biology and therapeutic applications for obesity and cardiometabolic disease. *Nat. Rev. Endocrinol.* **17**, 592–607 (2021).
72. Matthews, G. A. & Tye, K. M. Neural mechanisms of social homeostasis. *Ann. N. Y. Acad. Sci.* **1457**, 5–25 (2019).
73. Elovinio, M. et al. Contribution of risk factors to excess mortality in isolated and lonely individuals: an analysis of data from the UK Biobank cohort study. *Lancet Public Health* **2**, e260–e266 (2017).
74. Matthews, T. et al. Social isolation, loneliness and depression in young adulthood: a behavioural genetic analysis. *Soc. Psychiatry Psychiatr. Epidemiol.* **51**, 339–348 (2016).
75. Uhlen, M. et al. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
76. Morrison, J., Knoblauch, N., Marcus, J. H., Stephens, M. & He, X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat. Genet.* **52**, 740–747 (2020).
77. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
78. Russell, D., Peplau, L. A. & Cutrona, C. E. The revised UCLA loneliness scale: concurrent and discriminant validity evidence. *J. Pers. Soc. Psychol.* **39**, 472–480 (1980).
79. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* <https://doi.org/10.1038/s41586-023-06592-6> (2023).
80. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
81. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, s13742-015-0047-8 (2015).
82. Miller, K. L. et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).
83. Alfaro-Almagro, F. et al. Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* **166**, 400–424 (2018).
84. Desikan, R. S. et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).
85. Fischl, B. et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
86. Julkunen, H. et al. Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nat. Commun.* **14**, 604 (2023).
87. Ritchie, S. C. et al. Quality control and removal of technical variation of NMR metabolic biomarker data in ~120,000 UK Biobank participants. *Sci. Data* **10**, 64 (2023).

88. Mutz, J., Roscoe, C. J. & Lewis, C. M. Exploring health in the UK Biobank: associations with sociodemographic characteristics, psychosocial factors, lifestyle and environmental exposures. *BMC Med.* **19**, 240 (2021).
89. Spitzer, R. L., Kroenke, K., Williams, J. B. W. & and the Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care Study. *JAMA* **282**, 1737–1744 (1999).
90. Craig, C. L. et al. International physical activity questionnaire: 12-country reliability and validity. *Med. Sci. Sports Exerc.* **35**, 1381–1395 (2003).
91. IPAQ Research Committee. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)—short and long forms. *IPAQ* <https://sites.google.com/view/ipaq/score> (2005).
92. Ainsworth, B. E. et al. 2011 compendium of physical activities: a second update of codes and MET values. *Med. Sci. Sports Exerc.* **43**, 1575–1581 (2011).
93. Schlosser, P. et al. Netboost: boosting-supported network analysis improves high-dimensional omics prediction in acute myeloid leukemia and Huntington’s disease. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **18**, 2635–2648 (2021).
94. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
95. Walker, K. A. et al. Proteomics analysis of plasma from middle-aged adults identifies protein markers of dementia risk in later life. *Sci. Transl. Med.* **15**, eadf5681 (2023).
96. Cline, M. S. et al. Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* **2**, 2366–2382 (2007).
97. Kolberg, L. et al. g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res.* **51**, W207–W212 (2023).
98. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
99. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).
100. Griss, J. et al. ReactomeGSA—efficient multi-omics comparative pathway analysis. *Mol. Cell. Proteomics* **19**, 2115–2124 (2020).
101. Slenter, D. N. et al. WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res.* **46**, D661–D667 (2018).
102. Jiang, L. et al. A resource-efficient tool for mixed model association analysis of large-scale data. *Nat. Genet.* **51**, 1749–1755 (2019).
103. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
104. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
105. Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
106. Cochran, W. G. The combination of estimates from different experiments. *Biometrics* **10**, 101–129 (1954).
107. Hemani, G. et al. The MR-Base platform supports systematic causal inference across the human genome. *eLife* **7**, e34408 (2018).
108. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
109. Imai, K., Keele, L. & Tingley, D. A general approach to causal mediation analysis. *Psychol. Methods* **15**, 309–334 (2010).
110. Aalen, O. in *Mathematical Statistics and Probability Theory* (eds Klonecki, W., Kozek, A. & Rosiński, J.) 1–25 (Springer, 1980); https://doi.org/10.1007/978-1-4615-7397-5_1
111. Shen, C. GWAS summary statistics. *figshare* <https://doi.org/10.6084/m9.figshare.27261132.v1> (2024).

Acknowledgements

C.S. is supported by grants from the National Natural Sciences Foundation of China (no. 82101617) and the China Postdoctoral Science Foundation (no. 2022M710765). W.C. is supported by grants from the National Natural Sciences Foundation of China (no. 82071997) and the Shanghai Rising-Star Program (no. 21QA1408700). J.F. is supported by National Key R&D Program of China (no. 2018YFC1312904 and no. 2019YFA0709502), the Shanghai Municipal Science and Technology Major Project (no. 2018SHZDXZ01), the 111 Project (no. B18015), Shanghai Center for Brain Science and Brain-Inspired Technology and the Zhangjiang Lab. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

C.S., W.C. and J.F. conceived and designed the study. C.S. performed the data analysis with support from R.Z. All authors contributed to interpretation of the results. C.S. drafted the paper. B.J.S., C.S., J.Y. and W.C. revised the paper. C.S. and R.Z. contributed to the visualization. C.S., W.C. and J.F. had full access to the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. All authors read and approved the final manuscript.

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/s41562-024-02078-1>.

Correspondence and requests for materials should be addressed to Barbara J. Sahakian, Wei Cheng or Jianfeng Feng.

Peer review information *Nature Human Behaviour* thanks Fangyu Liu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Jianfeng Feng

Wei Cheng

Barbara Sahakian

Corresponding author(s):

Last updated by author(s): Oct 23, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	No software was used.
Data analysis	R version 4.2.0 was primarily used for the analyses in this study. Related R packages include: Protein co-expression analysis was performed using 'Netboost' (v2.4.1) (https://www.bioconductor.org/packages/release/bioc/html/netboost.html); Protein imputation for network analysis was performed using 'impute' (v1.70.0) (https://bioconductor.org/packages/release/bioc/html/impute.html); Functional enrichment was performed using 'gprofiler2' (v0.2.3) (https://cran.r-project.org/web/packages/gprofiler2/index.html); Protein GWASs were performed through GCTA (v1.94.1) (https://yanglab.westlake.edu.cn/software/gcta); Mendelian randomization analysis was performed using 'TwoSampleMR' (v0.5.7) (https://mrcieu.github.io/TwoSampleMR/), 'gsmr2' (v1.1.1) (https://yanglab.westlake.edu.cn/software/gsmr/), and 'cause' (v1.2.0) (https://github.com/jean997/cause); Colocalization was performed through 'coloc' (v5.2.3) (https://chr1swallace.github.io/coloc/); Preprocessing of NMR-metabolic data was performed using 'ukbnmr' (v2.2) (https://cran.r-project.org/web/packages/ukbnmr/index.html). Social isolation and loneliness GWASs were performed through PLINK 2.0 (https://www.cog-genomics.org/plink/2.0/). PPI was performed using STRING (v12.0) (https://string-db.org/cgi/input?sessionId=besj5Fp9GYc2&input_page_show_search=on). PPI network visualization was conducted by Cytoscape (v3.10.0) (https://cytoscape.org/). T1-weighted images were processed with Freesurfer (v6.0.0) (https://surfer.nmr.mgh.harvard.edu/). Custom scripts for the analyses have been made available through the following GitHub repository: https://github.com/chunshen617/Proteomics_loneliness .

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data used in the present study are available from the UK Biobank (<https://www.ukbiobank.ac.uk>) with restrictions applied. Data were used under licence and are thus not publicly available. Details regarding registration for data access can be found at <http://www.ukbiobank.ac.uk/register-apply/>. The data used in this study were accessed from the UK Biobank under the application number 19542. GWAS summary statistics used can be found at the Figshare website (https://figshare.com/projects/GWAS_summary_data/224229) (ref. 111). European ancestry reference data from the 1000 Genomes Project can be found via <https://github.com/getian107/PRScsx?tab=readme-ov-file>.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Both male and female subjects from the UK Biobank study were included. Sex (Data-Field 31) was obtained from the central registry at recruitment, but in some instances, it was updated by the participant. Consequently, this field may contain a blend of the sex recorded by the NHS for the participant and self-reported sex. Summary statistics on sex distributions were reported in Table 1. All statistical models were adjusted for sex, and sensitivity analyses including a sex interaction term and sex subgroup analysis were also conducted.

Reporting on race, ethnicity, or other socially relevant groupings

Table 1 presents the baseline characteristics of study participants, encompassing age, sex, ethnicity (white, mixed, Asian, black, and other), education level, and household income. Covariates were chosen based on literature and data availability. Two types of models, each with distinct covariates, were investigated in this study. The simple model comprised age, sex, site, technical factors, and the first 20 genetic PCs. Additionally, a fully adjusted model was explored, incorporating ethnicity, education level, household income, smoking, alcohol consumption, and BMI as covariates. The fully adjusted model was utilized for all primary analyses.

Population characteristics

Our primary study population included 42,062 participants (56.4 ± 8.2 years and 52.3% female) from the UK Biobank, who had quality-controlled proteomic data and complete behavioral data including social isolation, loneliness, and all covariates. A flow chart of participant selection is shown in Supplementary Fig. 1. Among these, 3,905 (9.3%) reported as being socially isolated, and 2,689 (6.4%) felt lonely. Detailed demographic characteristics, stratified by social isolation and loneliness, are presented in Table 1. During a median (SD) follow-up of 13.7 (2.1) years (ended on November 30, 2022 or death), 2,695 participants developed CVD, 892 developed all-cause dementia, 1,703 developed T2D, 1,521 developed depression, 983 developed stroke, and 4,255 passed away.

Recruitment

The UK Biobank is a population-based cohort that involves over 500,000 individuals aged 40–69 years recruited from 22 centers across the UK between 2006 and 2010. Previous investigation have demonstrated a healthy volunteer bias in the UK Biobank.

Ethics oversight

All participants in the UK Biobank provided informed consent, and ethical approval was obtained from the National Information Governance Board for Health and Social Care and the North West Multi-Centre Research Ethics Committee (ref: 11/NW/0382).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

No statistical methods were used to predetermine sample sizes and all currently available sample in the UK Biobank were used. Our primary study population included 42,062 participants (56.4 ± 8.2 years and 52.3% female) from the UK Biobank, who had quality-controlled proteomic data and complete behavioral data including social isolation, loneliness, and all covariates.

Data exclusions

In survival analyses, we excluded prevalent cases, which were defined as diagnoses occurring within the first three years of follow-up or self-reported cases at baseline. This exclusion was implemented to reduce the risk of potential reverse-causation bias. Genome-wide association studies (GWASs) on plasma protein abundance were conducted in Caucasian participants with qualified genotyping

data from the UK Biobank. To avoid over-fitting in two-sample MR, GWASs for social isolation (N=297,396) and loneliness (N=288,696) were performed using a distinct set of Caucasian participants, excluding those involved in the protein-related GWASs.

Replication

We replicated the primary analyses exclusively within the randomly selected subset (N=36,250). The proteins identified in this subset were consistent with those in all available samples, and the proteomic association patterns in these two populations were highly correlated (Supplementary Fig. 4). Additionally, to mitigate potential population stratification, we conducted PWAS specifically in Caucasians (N=35,697) and observed that the proteomic associative patterns were highly correlated with those found in the full sample (Supplementary Fig. 7). We also performed cross-validation by randomly splitting the UK Biobank samples 100 times. Our results showed that most of the significant proteins identified using the full sample retained significance in at least one of the two split samples, and proteomic associative patterns for social isolation and loneliness between the two split samples exhibited medium to large correlations (Supplementary Fig. 14).

Randomization

The primary analyses were controlled for age, sex, site, technical factors, ethnicity, education level, household income, smoking, alcohol consumption, BMI, and the first 20 genetic PCs. Additionally, two types of subgroup analyses were undertaken to examine potential sex (male vs female) and age (<60 vs >=60 years) differences.

Blinding

Blinding was not applicable to this study as this study is observational.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

The UK Biobank plasma samples were analysed using the Olink Explore 3072 proximity extension assay platform, which is based upon an in-solution binding of two polyclonal antibody pools to a target protein and subsequent hybridization and enrichment of two unique single-stranded DNA probes to create a double stranded barcode unique for the antigen. The platform consists of 2,941 immunoassays targeting 2,925 proteins. Each assay is based on a pair of polyclonal antibodies. The antibodies bind to different sites on the target protein and are labelled with single-stranded complementary oligonucleotides. If matching pairs of antibodies bind to the protein, the attached oligonucleotides hybridize, and are then measured using next-generation sequencing. Olink Explore 3072 consists of 8 panels of 384 assays analysed by next-generation sequencing.

Validation

Quality control protocol was implemented by the UK Biobank Pharma Proteomics Project, and was developed and approved by scientists across the thirteen participating biopharmaceutical companies, including Amgen, Alnylam, AstraZeneca, Biogen, Bristol Myers Squibb, Calico, Genentech, GlaxoSmithKline, Janssen (Johnson & Johnson), Novo Nordisk, Pfizer, Regeneron, and Takeda. More detailed information can be found in biobank.ndph.ox.ac.uk/ukb/docs/PPP_Phase_1_QC_dataset_companion_doc.pdf.

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.

Magnetic resonance imaging

Experimental design

Design type	Structural MRI
Design specifications	The UK Biobank designed the imaging acquisition protocols including 6 modalities, covering structural, diffusion and functional imaging. The collection order is T1-weighted structural image, resting-state functional MRI, task functional MRI, T2-weighted FLAIR structural image, diffusion MRI and susceptibility-weighted imaging. The T1-weighted structural image was acquired using straight sagittal orientation for 5 minutes.
Behavioral performance measures	We used T1-weighted structural imaging, during which participants were not required to perform any tasks.

Acquisition

Imaging type(s)	T1-weighted structural imaging
Field strength	3T
Sequence & imaging parameters	The EPI-based acquisitions utilize simultaneous multi-slice (multiband) acceleration. UK Biobank uses pulse sequences and reconstruction code from the Center for Magnetic Resonance Research (CMRR), University of Minnesota https://www.cmrr.umn.edu/multiband . The resolution is 1x1x1 mm and the field of view is 208x256x256 matrix. Straight sagittal orientation is used. TR and TE are 2000ms and 2.01ms respectively. The flip angle is 8 deg. Detailed sequence and imaging parameters are openly available here: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/brain_mri.pdf .
Area of acquisition	Whole brain
Diffusion MRI	<input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used

Preprocessing

Preprocessing software	Imaging-derived phenotypes (IDPs) generated through an image-processing pipeline developed and run on behalf of the UK Biobank were used in this study (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf). T1-weighted images were processed using FreeSurfer. Where available, T2_FLAIR images were used in conjunction with T1-weighted images to achieve more accurate cortical modeling than with T1 alone. Surface atlases were employed to extract IDPs related to surface area, volume, and mean cortical thickness of standard atlas regions. The Qoala-T approach was used to assess the quality of FreeSurfer outputs, supported by manual checks for outputs near the quality threshold. Any FreeSurfer outputs failing quality control were excluded from the IDPs. Surface volumes extracted using the Desikan-Killiany atlas for 68 cortical regions and the aseg atlas for 16 subcortical regions were used in this study. The full processing pipeline is openly available here: https://doi.org/10.1016/j.neuroimage.2017.10.034 and https://www.nature.com/articles/nrn.4393 .
Normalization	Spatial normalization was performed whenever applicable as described in detail in Miller et al. Nature Neuroscience (2016) and Alfaro-Almagro et al. Neuroimage (2018).
Normalization template	1 mm resolution version of MNI152 template
Noise and artifact removal	Noise and artifact removal were described in detail in Miller et al. Nature Neuroscience (2016) and Alfaro-Almagro et al. Neuroimage (2018).
Volume censoring	No volume censoring performed on this data

Statistical modeling & inference

Model type and settings	Linear regression model was used to investigate the associations between proteins and IDPs. Covariates include age, sex, imaging collection site, batch, time gap between blood collection and protein measurement, ethnicity, education level, household income, smoking, alcohol consumption, BMI, the first 20 genetic PCs, ICV, and time gap between baseline and imaging collection.
Effect(s) tested	T-tests were used to derive the two-sided p value.
Specify type of analysis:	<input type="checkbox"/> Whole brain <input checked="" type="checkbox"/> ROI-based <input type="checkbox"/> Both
Anatomical location(s)	Surface volumes extracted using the Desikan-Killiany atlas for 68 cortical regions and the aseg atlas for 16 subcortical regions were used in this study.
Statistic type for inference	No whole brain voxel-wised or cluster-based analyses involved in this study.

(See [Eklund et al. 2016](#))

Correction

Both Bonferroni ($P < 0.05/(84*5)$) and FDR correction ($q < 0.05$) using the Benjamini-Hochberg procedure results are provided. For FDR correction, all tests are considered simultaneously.

Models & analysis

n/a	Involved in the study
<input checked="" type="checkbox"/>	Functional and/or effective connectivity
<input checked="" type="checkbox"/>	Graph analysis
<input checked="" type="checkbox"/>	Multivariate modeling or predictive analysis