

# Social disadvantage accelerates aging

Received: 28 May 2024

Accepted: 4 February 2025

Published online: 14 March 2025

Check for updates

Mika Kivimäki<sup>1,2</sup>✉, Jaana Pentti<sup>2,3</sup>, Philipp Frank<sup>1,2</sup>, Fangyu Liu<sup>4</sup>, Acer Blake<sup>5,6</sup>, Solja T. Nyberg<sup>2</sup>, Jussi Vahtera<sup>3</sup>, Archana Singh-Manoux<sup>1,7</sup>, Tony Wyss-Coray<sup>8,9,10</sup>, Keenan A. Walker<sup>4</sup>, Linda Partridge<sup>5</sup> & Joni V. Lindbohm<sup>1,2,11</sup>

Social disadvantage, like advanced age, is a risk factor for a broad range of health conditions; however, whether it influences the aging process remains unclear. Here, using a multicohort approach, we investigated the associations of social disadvantage with age-related plasma proteins and age-related diseases. We found proteomic signatures of accelerated immune aging and 14 specific age-related proteins linked to social disadvantage during both early and later life. Individuals experiencing social disadvantage had an increased risk of 66 age-related diseases, with up to 39% of these associations mediated by the 14 age-related proteins (for example, DNAJB9, F2, HSPA1A, BGN). The main enriched pathway involved the upregulation of the pro-inflammatory regulator NF- $\kappa$ B24 and its downstream factor interleukin-8. Our findings support the hypothesis that social disadvantage throughout the life course may accelerate aging, a biological mechanism that could explain why social stratification plays such a fundamental role in determining human health.

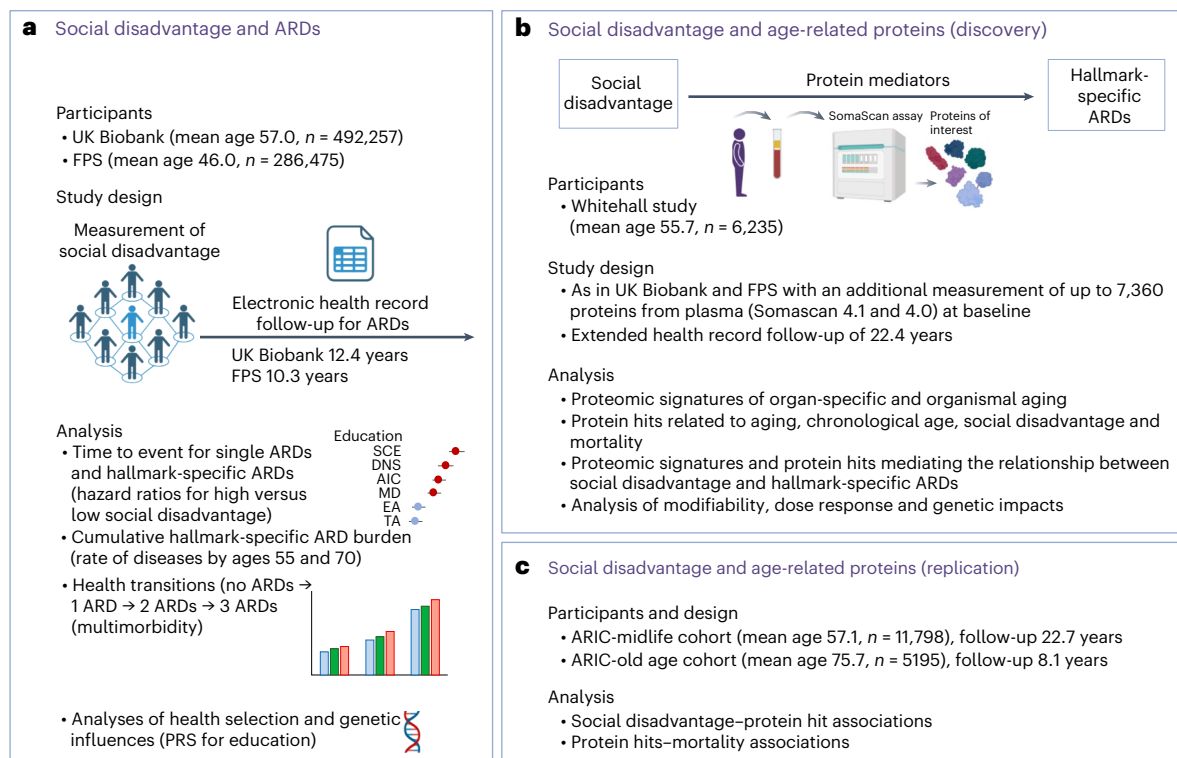
Most human health conditions are socially patterned<sup>1–3</sup>. Social disadvantage is a known risk factor for mental and behavioral disorders, poor glucose and insulin metabolism, and impaired immune function, as well as diseases of the liver, kidney, and vascular and respiratory systems, site-specific cancers and dementia<sup>3</sup>. Such a breadth of health effects is rarely observed for a single risk factor, except age. As with social disadvantage, the progressive cellular damage associated with aging increases susceptibility to most diseases<sup>4–6</sup>.

Parallels between the effects of social disadvantage and biological aging suggest that social disadvantage may accelerate the aging process<sup>7–9</sup>. Aging is associated with declining metabolic function, including increased insulin resistance, impaired hepatic gluconeogenesis, dysregulated adipose lipogenesis and atherosclerosis<sup>10,11</sup>. Many components of the immune system also deteriorate with age. These changes include an increase in levels of circulating pro-inflammatory

mediators and a downregulation of immune responsiveness<sup>12,13</sup>. Notably, the same changes are characteristic of the effects of social disadvantage<sup>12,14–17</sup>. In addition, there is evidence linking social disadvantage to biological aging, as measured by epigenetic aging clocks<sup>7–9</sup>.

If social disadvantage accelerates the aging process, it should be linked to one or more of the nine original hallmarks of cellular aging<sup>4,5,18</sup>. These include four primary contributors to age-related cellular damage—genomic instability (GI), telomere attrition (TA), epigenetic alterations (EA) and loss of proteostasis (LOP, the breakdown of the cellular mechanisms that regulate the proper synthesis, folding, trafficking and degradation of proteins). In addition, they comprise five compensatory or integrative hallmarks—deregulated nutrient sensing (DNS), mitochondrial dysfunction (MD), cellular senescence (CS), altered intercellular communication (AIC) and stem cell exhaustion (SCE)<sup>5</sup>. Although not diseases themselves, these

<sup>1</sup>Brain Sciences, University College London, London, UK. <sup>2</sup>Clinicum, University of Helsinki, Helsinki, Finland. <sup>3</sup>Department of Public Health and Centre for Population Health Research, University of Turku, Turku University Hospital, Turku, Finland. <sup>4</sup>Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA. <sup>5</sup>Institute of Healthy Ageing, GEE, University College London, London, UK. <sup>6</sup>MPLS (Mathematical, Physical and Life Sciences) Division, Oxford University, Oxford, UK. <sup>7</sup>Inserm U1153, Epidemiology of Ageing and Neurodegenerative Diseases, Université Paris Cité, Paris, France. <sup>8</sup>Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA. <sup>9</sup>Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, USA. <sup>10</sup>The Phil and Penny Knight Initiative for Brain Resilience, Stanford University, Stanford, CA, USA. <sup>11</sup>Broad Institute of MIT and Harvard University, Cambridge, MA, USA. ✉e-mail: [m.kivimaki@ucl.ac.uk](mailto:m.kivimaki@ucl.ac.uk)



**Fig. 1 | Overview of the study design.** **a**, Summary of study characteristics and analyses from the UK Biobank and the FPS. Indicators of social disadvantage were education and adulthood SES, measured by residential neighborhood deprivation and occupational position. In individuals with social disadvantage, a higher risk of ARDs was observed across all nine hallmarks of aging. This was evident for the onset of the first ARD, cumulative ARD burden and ARD multimorbidity. By contrast, there was limited evidence for the reverse direction of the relationship, in which ARDs lead to social disadvantage, or for genetic factors explaining the observed associations. **b**, The characteristics and analyses of the Whitehall study, including proteomic data, are summarized. With the exception of immune function and kidney aging, proteomic signatures of organ-specific aging were not strongly associated with social disadvantage. By contrast,

of the 1,040 hallmark-related proteins associated with chronological age at proteome-wide significance, 14 were consistently linked to indicators of social disadvantage and hallmark-specific ARDs, partially mediating this association. Additional analyses supported the modifiability of these proteins in relation to changes in social disadvantage and a dose–response association indicating accumulated risk. **c**, Characteristics and analyses of the two cohorts of the ARIC study, including data on 11 of the 14 proteins, are summarized. The associations between social disadvantage and protein levels replicated the findings from the Whitehall study. Similarly, the longitudinal associations between these proteins and all-cause mortality, with both 23-year and shorter 8-year follow-ups, were consistent with those observed in Whitehall. Figure created with [BioRender.com](#).

hallmarks contribute to age-related conditions, such as SCE increasing cardiovascular risk, CS heightening susceptibility to infections and GI being linked to cancer<sup>4,5,18</sup>.

To better understand the links between social disadvantage and biological aging, we examined whether social disadvantage is associated with 83 age-related diseases (ARDs) and over 3,900 age-related circulating proteins, including those linked to accelerated organ-specific and organismal aging<sup>19,20</sup> (Fig. 1). Our findings support the concept that social disadvantage may accelerate age-related changes, aligning with the social causation hypothesis. While reverse causality—in which ARDs contribute to social disadvantage—and genetic predisposition influencing both the likelihood of social disadvantage and the development of ARDs were also observed, their roles appear more limited.

## Results

This study draws on four prospective observational studies, comprising five cohorts analyzed sequentially. Two large-scale studies, the UK Biobank and the Finnish Public Sector Study (FPS), were used to test the social causation hypothesis in relation to ARDs associated with hallmarks of aging (Fig. 1a). We tested the alternative health-related selection hypothesis using the FPS, with its repeated measures of social disadvantage, to examine whether ARDs increase the risk of social disadvantage. In addition, genetic data from the UK Biobank enabled an investigation into the genetic contributions to these associations<sup>21</sup>.

To identify age-related protein mediators of the associations between social disadvantage and hallmark-specific ARDs, we used proteomic data from the UK Whitehall study (Fig. 1b) and replicated the main findings in the US Atherosclerosis Risk in Communities (ARIC) study, including both midlife and older-age cohorts (Fig. 1c).

In these four studies, social disadvantage was measured in both early and later life (Table 1, Extended Data Figs. 1 and 2, and Supplementary Table 1). Education, typically spanning ages 5–7 to ages 15–25, was assessed in all cohorts as a marker of social standing in early life and across the life course. In the Whitehall study, father's social class—representing influences from birth until independent living—served as an additional early-life indicator. In all cohorts, neighborhood deprivation was used as a measure of adult socioeconomic status (SES). Other indicators of adult SES included occupational position in the FPS and Whitehall cohorts, and family income in the ARIC cohorts.

### Social disadvantage and risk of ARDs

As an initial step, we selected ARDs associated with hallmarks of aging. These included a total of 83 diseases linked to one or more hallmarks of aging, based on the taxonomy put forward in ref. 4 (Supplementary Table 2). Support for this taxonomy comes from multiple sources. Analyses of electronic health records from general practice and hospitalizations identified more than 200 diseases with incidence rates increasing with chronological age<sup>6,22</sup>. Researchers linked a subset of

**Table 1 | Baseline characteristics of participants by cohort**

Characteristics	UK Biobank		FPS		Whitehall		ARIC midlife		ARIC old age	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Number of participants	492,257	100	286,475	100	6,544	100	11,798	100	5,195	100
Age (years), mean (s.d.)	57.0 (8.1)		46.0 (7.4)		55.7 (6.0)		57.1 (5.7)		75.7 (5.2)	
Sex										
Men	224,286	45.6	76,902	26.8	4,641	70.9	5,242	44.4	2,970	57.2
Women	267,971	54.4	209,573	73.2	1,903	29.1	6,556	55.6	2,225	42.8
Ethnicity										
White	464,816	94.4			5,982	91.4	8,987	76.2	4,200	80.8
Non-White	25,768	5.2			562	8.6	2,733	23.2	966	18.6
Missing	1,673	0.3			0	0.0	78	0.7	29	0.6
PRS for education										
Mean (s.d.)	0.0 (1.0)	97.7			0.0 (1.0)	48.9				
Missing	11,550	2.3			3,187	51.1				
<b>Early-life social indicator</b>										
Education										
Primary	85,258	17.3	34,693	12.1	2,336	35.7	2,564	21.7	700	13.5
Secondary	245,883	50.0	96,917	33.8	1,736	26.5	4,948	41.9	2,192	42.2
Tertiary	161,116	32.7	154,865	54.1	2,339	35.7	4,269	36.2	2,295	44.2
Missing	0	0.0	0	0.0	133	2.0	17	0.1	8	0.2
<b>Adult SES indicators</b>										
Neighborhood deprivation										
Low	124,225	25.2	89,606	31.3	1,575	24.1	4,139	35.1	1,735	33.4
Intermediate	245,783	49.9	148,197	51.7	3,167	48.4	3,538	30.0	1,735	33.4
High	121,637	24.7	37,970	13.3	1,513	23.1	3,787	32.1	1,712	33
Missing	612	0.1	10,702	3.7	289	4.4	334	2.8	13	0.3
Occupational status or income										
Low			79,431	27.7	863	13.2	3,987	33.8	1,870	36.0
Intermediate			74,090	25.9	2,801	42.8	4,283	36.3	1,599	30.8
High			120,194	42.0	2,806	42.9	2,882	24.4	1,238	23.8
Missing			12,760	4.5	74	1.1	646	5.5	488	9.4

For additional indicators of social disadvantage, see Supplementary Table 1.

these ARDs to specific hallmarks of aging using several approaches: mining 1.85 million PubMed abstracts on human aging, identifying shared genes in the genome-wide association study catalog, conducting gene set enrichment analysis and analyzing disease co-occurrence networks within each hallmark<sup>4</sup>.

We confirmed the co-occurrence of ARDs within each hallmark in 492,257 participants from the UK Biobank study<sup>23</sup>. The presence of one ARD increased the risk of developing another ARD related to the same hallmark, with clustering coefficients ranging from 0.76 for LOP-specific ARDs to 0.92 for SCE-specific ARDs. These findings corroborated the hallmark-specific clustering of ARDs (Extended Data Figs. 3 and 4)<sup>23</sup>.

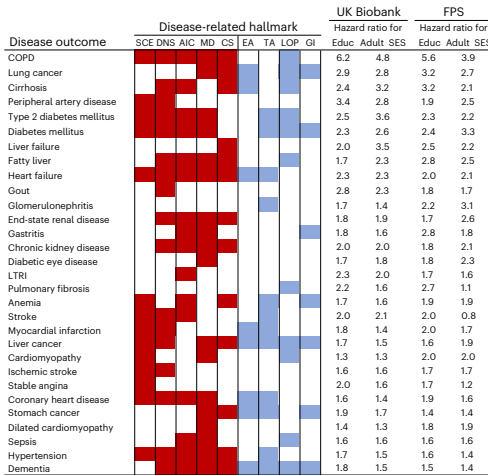
In time-to-event analyses of UK Biobank and FPS participants without these ARDs at baseline (*n* ranging from 477,325 to 492,294 in the UK Biobank and from 278,272 to 286,471 in the FPS, depending on the social disadvantage indicator and ARD), social disadvantage—indicated by education and adult SES (neighborhood deprivation)—was associated with a higher risk of developing ARDs. In the UK Biobank, the age-, sex- and ethnicity-adjusted hazard ratio for developing any ARD was 1.31 (95% confidence interval (CI) 1.29–1.33) for individuals with low compared with high education. For individuals with high versus

low adult SES, the hazard ratio was 1.21 (95% CI 1.20–1.23). In the FPS, the corresponding hazard ratios were 1.28 (95% CI 1.25–1.31) and 1.23 (95% CI 1.20–1.27), respectively.

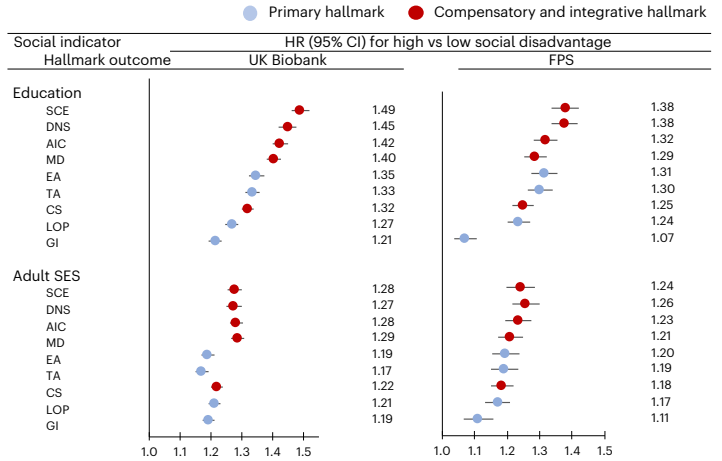
Overall, social disadvantage was associated with an increased risk of 66 specific ARDs (Supplementary Table 3). The 30 ARDs most strongly associated with social disadvantage were multi-hallmark diseases, as classified by the taxonomy in ref. 4 (Fig. 2a). These included liver, metabolic, circulatory, respiratory, infectious and neurodegenerative diseases. By contrast, social disadvantage was less strongly associated with diseases linked to a single primary or compensatory and integrative hallmark of aging (Extended Data Fig. 5).

In separate time-to-event analyses of the first-onset ARD within each of the nine hallmarks, individuals with social disadvantage had a higher risk of ARDs across all hallmarks compared with those without social disadvantage (Extended Data Fig. 6 and Supplementary Table 4). In both the UK Biobank and FPS, these associations were generally stronger for ARDs linked to compensatory and integrative hallmarks than for those associated with primary hallmarks (Fig. 2b). This was particularly evident for SCE-, DNS-, AIC- and MD-specific ARDs, while the weakest association was observed for GI-specific ARDs. These findings were robust, as similar patterns were observed in analyses using

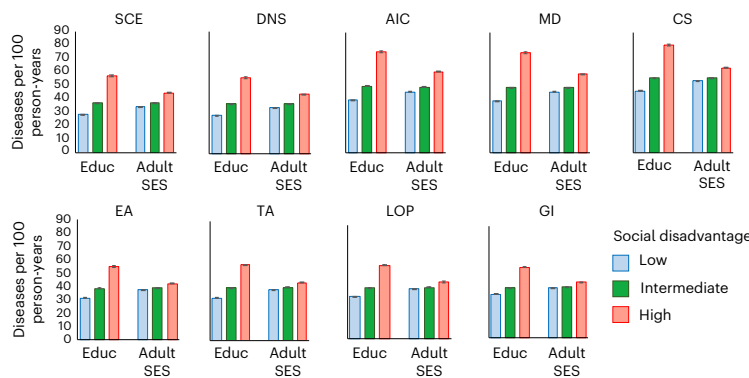
**a** Outcome: top 30 ARDs



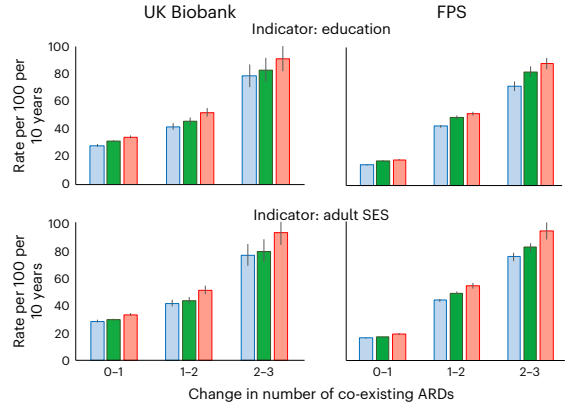
**b** Outcome: first hallmark-specific ARD



**c** Outcome: cumulative burden of hallmark-specific ARDs



**d** Health transitions from healthy to ARD multimorbidity



**Fig. 2 | Social disadvantage and risk of diseases associated with hallmarks of aging.** These analyses examine the social causation hypothesis using two-sided tests without adjustments for multiple comparisons (for full results, see Supplementary Tables 3–12). The whiskers represent 95% CIs. **a**, The numbers indicate hazard ratios from Cox proportional hazards models comparing high versus low social disadvantage at baseline (education and adulthood SES), adjusted for age, sex and ethnicity, for a single ARD risk at follow-up in the UK Biobank and FPS (sample size 481,197–492,257 in the UK Biobank and 275,157–285,830 in the FPS, depending on the social disadvantage indicator and ARD). All HRs are statistically significant ( $P < 0.05$ ). Only the 30 strongest associations are shown. **b**, The forest plot shows hazard ratios from Cox models for developing hallmark-specific ARDs, comparing participants with high versus low social disadvantage in a population free of these diseases at baseline. All HRs

are statistically significant ( $P < 0.05$ ). In the UK Biobank, sample sizes ranged from 430,373 to 460,980, with the corresponding FPS sample sizes between 257,073 and 281,348. **c**, The bars show the number and 95% CIs of hallmark-specific ARDs per 100 person-years by age 70 in the UK Biobank, stratified by social disadvantage level, estimated using Poisson regression (range of person-years 34,134,089–34,174,830). The results for FPS are available in Supplementary Table 7 (range of person-years 15,508,103–16,112,263). **d**, The bars represent ARD multimorbidity progression rates, from an ARD-free state to three co-occurring ARDs, stratified by social disadvantage levels, based on Poisson regression analysis in a population free of these diseases at baseline (range of person-years 4,959,308–4,965,499 in UK Biobank and 2,614,663–2,712,578 in FPS). HR, hazard ratio; Educ, education.

an alternative indicator of adult SES—occupational position—available for FPS participants and in both men and women, although the effects were stronger in men (Supplementary Tables 4 and 5).

Given the higher mortality rate among individuals with social disadvantage, we examined whether competing risks, such as deaths, might bias the associations between social disadvantage and the risk of hallmark-specific ARDs. However, no such evidence was found, as these associations persisted in competing risk analyses with hallmark-specific ARDs and death as outcomes (Supplementary Table 6).

To determine whether social disadvantage was associated with the total hallmark-specific ARD burden, including both first and subsequent ARDs, we calculated the average number of hallmark-specific ARDs per 100 person-years by age 70 in participants from the UK Biobank. The age-, sex- and ethnicity-adjusted rate was up to two times higher in individuals with greater social disadvantage than

in those with less social disadvantage (Fig. 2c and Supplementary Tables 7 and 8). A similar pattern was observed when estimated up to age 55 in the younger FPS cohort compared with the UK Biobank population (Supplementary Tables 7 and 8). These findings were consistent across both indicators of social disadvantage—education and adult SES—in both cohorts, as well as for occupational positions in FPS participants.

To examine each health transition separately, we conducted a stepwise analysis focusing on the development of multiple ARDs during the follow-up, starting from a baseline healthy state with no ARDs. Transition rates from 0 to 1, 1 to 2 and 2 to 3 ARDs increased by 1.5 to more than 2 times at each step, with the highest rates consistently observed in the socially disadvantaged group at every transition (Fig. 2d and Supplementary Tables 9 and 10).

Overall, these results support a social causation model, suggesting that social conditions in both early and later life influence the risk

of ARDs and contribute to health disparities across groups defined by educational background, residential neighborhoods and occupational status. In a model including both education and adult SES, each indicator showed an independent association with ARDs with the effect of education being stronger than that of neighborhood deprivation (Supplementary Table 11).

### Health-related selection into social disadvantage

We then examined the reverse causality direction, wherein ARDs are hypothesized to negatively impact subsequent levels of adult SES, a concept known as the health-related selection model. Among FPS participants whose adult SES changed between baseline and follow-up, those with an ARD at baseline were more likely to experience a downward shift in adult SES compared with participants without an ARD at baseline (Supplementary Table 12). However, the effect estimates—indicating a 1.1 to 1.2 times higher risk of declining SES over the next 5 years ( $n = 58,000$ ) or 10 years ( $n = 39,608$ ) for individuals with pre-existing ARD—were weaker than those supporting the social causation model. Thus, while both social causation and health-related selection processes are probably at play, social causation appears to have a stronger influence.

### Social disadvantage and age-related circulating proteins

ARDs represent outcomes of aging rather than the aging process itself. To better understand the aging process in relation to social disadvantage, beyond its outcomes, we examined the association between social disadvantage and blood-based proteomic signatures of aging. Our rationale for these analyses is threefold. First, aging is associated with measurable changes in the plasma proteome<sup>19,20</sup>. Circulating proteins regulate the aging process, while aging also modifies blood protein composition. Some proteins additionally serve as markers of aging, entering the plasma through leakage from damaged or dying cells. Second, ARDs and age-related proteins are empirically distinct constructs: the former are based on hospitalization diagnoses, and the latter on biomarker measurements, together providing a complementary assessment of aging from two independent information sources. Third, proteins reflect multiple age-related processes that precede ARDs. Analyzing them helps identify specific aging processes influenced by social disadvantage and the extent to which these processes mediate the associations of social disadvantage with individual aging outcomes.

The Whitehall study is one of the few long-term prospective cohort studies that provide extensive data on indicators of social disadvantage, link participants to electronic health records for assessing hallmark-specific ARDs and store plasma samples for biomarker measurements collected between 1997 and 1999, coinciding with the assessment of social disadvantage. In addition, consistent with findings from the UK Biobank and FPS, education and adult SES (including neighborhood deprivation and occupational position) were associated with hallmark-specific ARDs in the 22-year follow-up of the Whitehall study (Supplementary Table 13).

To analyze differences in age-related plasma proteins among 6,235 Whitehall participants with high and low levels of social disadvantage (mean age 56, 29% women), we assessed expression levels of up to 7,360 circulating proteins using the SomaLogic proteomic platform assays version 4.0 ( $n$  participants = 2,213) and version 4.1 ( $n = 4,022$ ). Using the Organage package for Python<sup>24</sup>, we estimated proteomic signatures for the biological age of each participant's organs relative to those of same-aged peers, that is, organ age gaps, as well as a non-organ-specific (organismal) age gap<sup>24</sup> (Supplementary Table 14).

After adjustment for age, sex and ethnicity, individuals with social disadvantage had higher age gaps in immune function, kidneys and lungs (Fig. 3a, Extended Data Fig. 7 and Supplementary Table 15). However, organ-specific age gaps mediated only a modest proportion of the associations between education and hallmark-specific ARDs

(0% to 6.2%), and between adult SES and hallmark-specific ARDs (5.4% to 13.9%) (Supplementary Table 16).

To identify additional and potentially stronger mediators linking social disadvantage to hallmark-specific ARDs, we performed a higher-resolution protein-level discovery analysis. This analysis focused on 1,040 proteins of the 7,360 SomaLogic platform, which were previously linked to one or more hallmarks of aging in the Human Proteomic Atlas ([www.proteinatlas.org](http://www.proteinatlas.org))<sup>4</sup> (Supplementary Table 17) and for which the expression levels in plasma were associated with chronological age at proteome-wide significance in the Whitehall study ( $P < 1.67 \times 10^{-6}$ ) (Supplementary Table 18). These proteins are derived from a source of information independent of the health records used to assess hallmark-specific ARDs, ensuring that their use in mediation analyses is methodologically sound.

We found 14 hallmark-related proteins that were consistently associated with age, indicators of social disadvantage and mortality (Fig. 3b and Supplementary Tables 19 and 20). Furthermore, these proteins were associated with six to nine hallmark-specific ARDs, depending on the protein under investigation (Fig. 3b and Supplementary Table 21). Sex-stratified analyses showed a similar overall pattern of results in both men and women (Supplementary Table 22).

### Protein mediators of social disadvantage-related disease risk

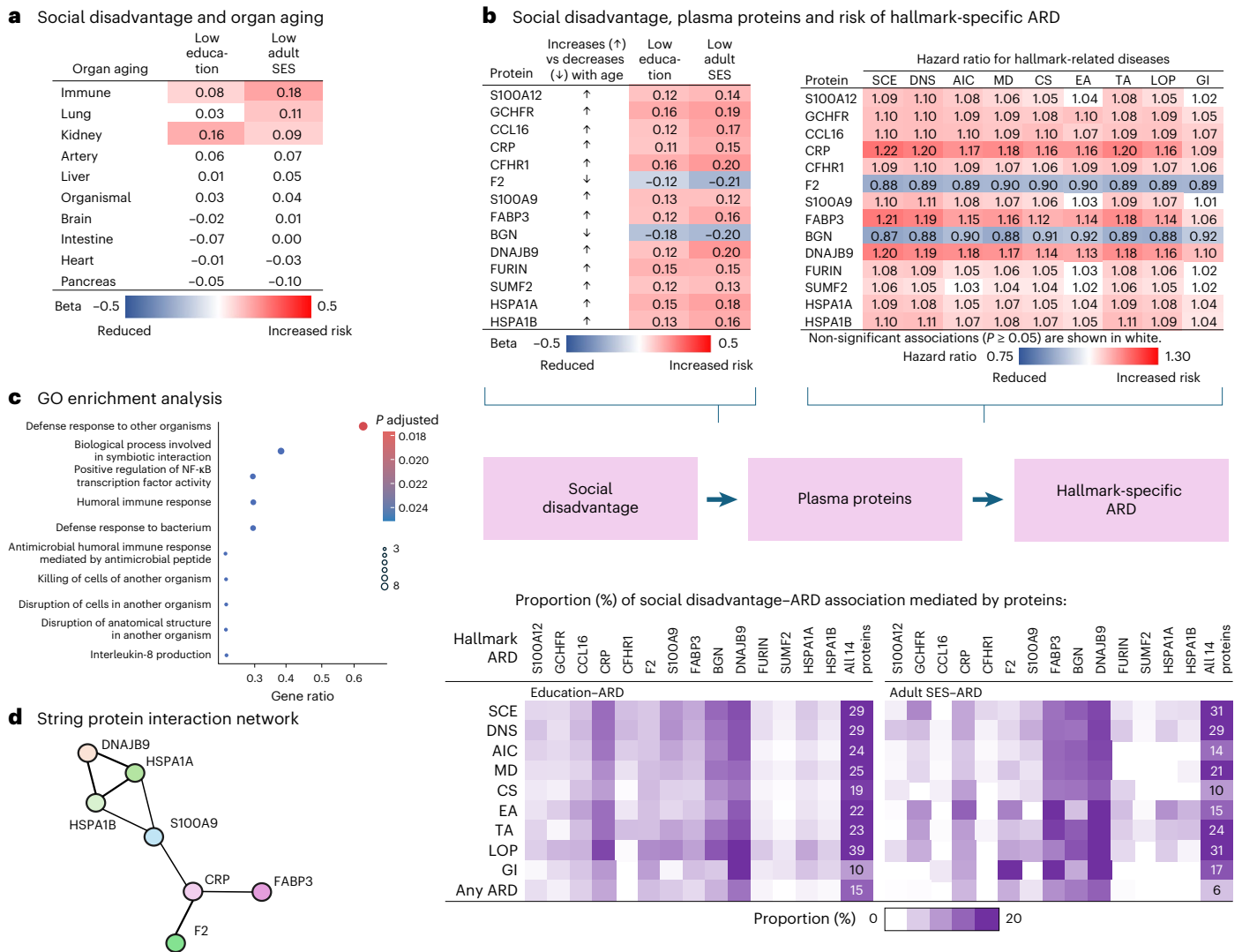
To quantify the contributions of the 14 proteins to the associations between social disadvantage and hallmark-specific ARDs, we conducted a series of mediation analyses. The 14 proteins collectively mediated 9.6% to 39.2% of the associations between education and hallmark-specific ARDs, and 9.9% to 30.6% of the associations between adult SES and hallmark-specific ARDs, depending on the hallmark-specific ARD group of interest (Fig. 3b and Supplementary Tables 23 and 24). For ARDs associated with SCE, DNS, MD, TA and LOP, the proportion of mediation exceeded 20.0%. Separate analyses identified DNAJB9, CRP and BGN as the strongest mediators for education, and DNAJB9, FABP3 and BGN (followed by CRP and F2) for adult SES, in their associations with hallmark-specific ARDs.

To facilitate biological interpretation, we performed pathway enrichment analysis for the 14 proteins, revealing their enrichment in processes activated in response to both internal and external inflammatory stimuli. The main enriched inflammatory pathway was the upregulation of the pro-inflammation regulator NF- $\kappa$ B (ref. 25) and its downstream factor interleukin-8, which is associated with cell senescence<sup>26</sup> (Fig. 3c and Extended Data Fig. 8). The related interaction network, formed by 7 of the 14 proteins, consisted of damage-associated molecular patterns (DAMPs) from the S100 protein family (S100A12 and S100A9), heat shock protein family (DNAJ9B, HSPA1A, HSPA1B) and C-reactive protein, which can recognize both DAMPs and pathogen-associated molecular patterns (PAMPs) (Fig. 3d).

### Protein concentrations reflect changes in social standing

We conducted additional analyses to assess whether the Whitehall data align with the social causation model regarding the associations between social disadvantage and the 14 proteins. First, if social disadvantage affects protein concentrations, the elimination of social disadvantage should relate to more favorable protein concentrations compared with persistent social disadvantage. Conversely, a shift from no social disadvantage to social disadvantage should be associated with less favorable protein concentrations (that is, concentrations related to a higher risk of ARDs) compared with a consistent absence of social disadvantage.

In the Whitehall study, this was supported in analyses combining early- and later-life indicators of social disadvantage (Fig. 4a and Supplementary Table 25). An improvement in social standing (that is, from low education to intermediate or high adult SES) was linked to more favorable protein concentrations compared with remaining in a low position (low education and low adult SES). A downward shift in



**Fig. 3 | Plasma proteins associated with social disadvantage and risk of hallmark-specific ARDs and mortality.** **a, b.** Findings from the Whitehall study; full results are in Supplementary Tables 15–24. **a.** The heat map shows results from multinomial logistic regression analyses examining the associations between social disadvantage and protein signatures of age gaps defined as the biological age of an individual’s organs or body relative to that of same-aged peers. The numbers represent beta coefficients for low versus high education and neighborhood deprivation, adjusted for age, sex and ethnicity. **b.** Of the 1,040 proteins with concentrations significantly associated with chronological age at the proteome-wide level (two-sided  $P < 1.67 \times 10^{-6}$ ), 14 proteins are highlighted for their consistent associations with age (linear regression analysis), social disadvantage indicators (cumulative logistic regression analysis) and mortality (Cox regression). The numbers in the upper heat

maps show beta coefficients and hazard ratios per s.d. increase in protein concentration, from linear regression for social disadvantage and Cox regression for ARDs, adjusted for age, sex and ethnicity. The lower heat map shows the proportion of the association between social disadvantage and hallmark-specific ARDs mediated by the 14 proteins, calculated using the inverse odds ratio-weighted method. **c.** GO enrichment analyses of the 14 proteins are shown. Rows show the GO terms, the dot sizes show the number of enriched proteins, the colors indicate the FDR-adjusted  $P$  value and the x axis shows the proportion of enriched proteins relative to all proteins associated with the GO term. **d.** String protein interaction network analysis indicates that 7 of the 14 proteins formed a protein interaction network. Only associations above a confidence score of 0.4 (standard medium confidence) are shown.

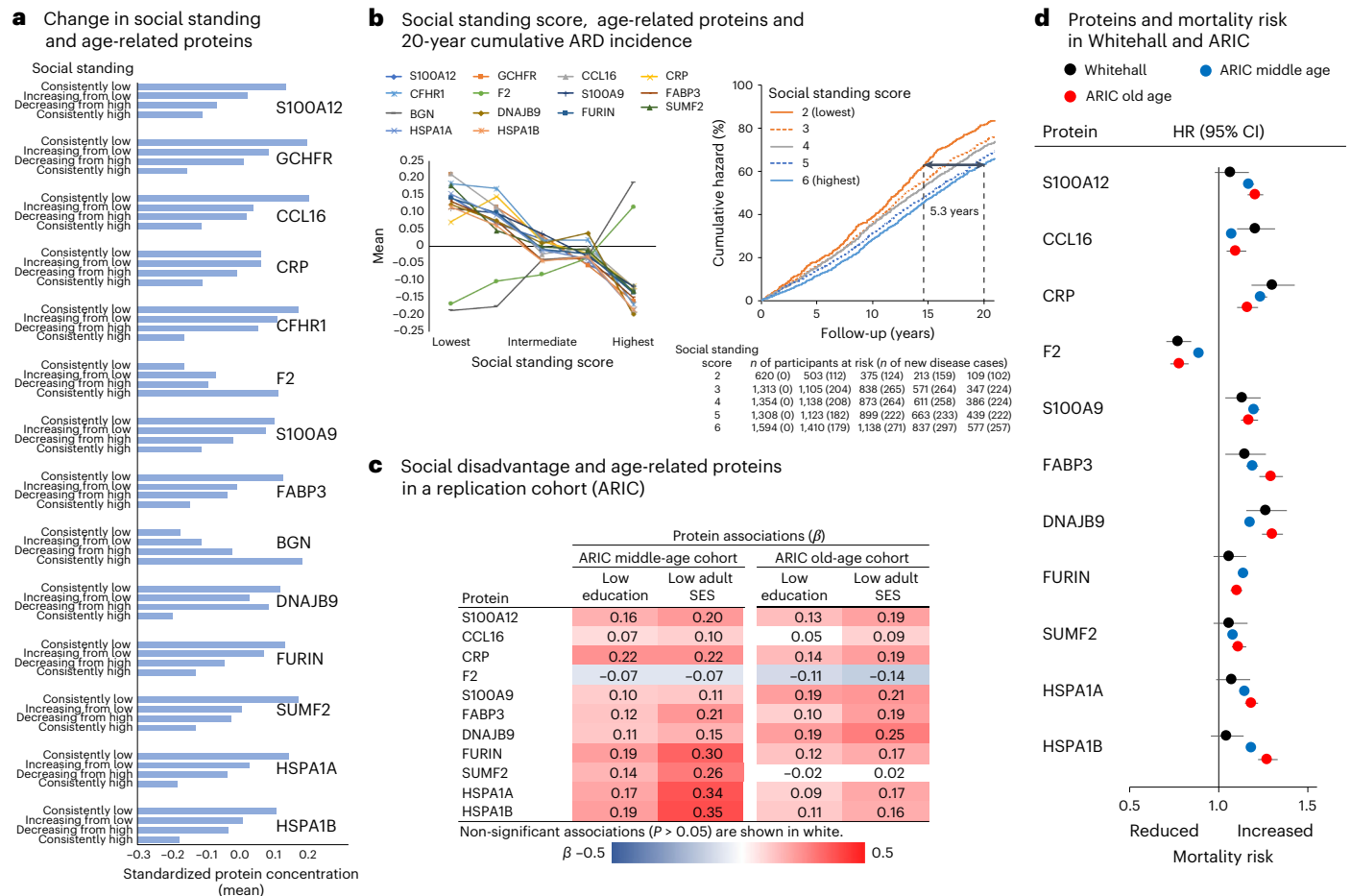
social standing (from high education to intermediate or low adult SES) was associated with less favorable protein concentrations compared with remaining in a high position (high education and high adult SES).

**Evidence for dose–response relationships**

To further investigate social causation, we examined whether a dose–response relationship existed between exposure to social disadvantage and the levels of age-related proteins. A graded association—in which greater social disadvantage corresponds to less favorable expression of age-related proteins—would support the social causation hypothesis. To quantify social disadvantage exposure for each Whitehall participant, we constructed a life course social standing score

(ranging from 2 to 6) by summing the 3-level education and the 3-level adult SES variables (low = 1, intermediate = 2, high = 3). The subsequent risk accumulation analysis supported the social causation model, showing a dose–response relationship across all 14 proteins, with higher social standing scores linked to more favorable protein concentrations (Fig. 4b and Supplementary Table 26).

To evaluate whether a similar dose–response relationship exists for aging outcomes, we examined the association between the life course social standing score and the risk of developing ARDs during the follow-up period. Consistent with the social causation model, we found a dose–response association with ARDs, similar to that observed with protein concentrations. Separation in the hazard curves between



**Fig. 4 | Supplementary and replication analyses for the associations between social disadvantage, protein concentrations, ARDs and mortality.** These results from analyses examining the social causation hypothesis are based on two-sided tests, without adjustment for multiple testing (for full results, see Supplementary Tables 25–30). **a**, The y axis shows age-, sex- and ethnicity-adjusted means of standardized protein concentrations by combinations of early- and later-life social disadvantage categories, measured by education and adult SES in the Whitehall study. Supporting modifiability, a reduction in social disadvantage was associated with more favorable protein concentrations in later life, compared with persistent social disadvantage. Conversely, the onset of social disadvantage was linked to less favorable protein concentrations, compared with remaining free from disadvantage. **b**, The left panel shows age-adjusted means of standardized protein concentrations across a life-course social standing score (range: 2–6), illustrating dose–response associations: a higher social standing corresponds to more favorable protein concentrations, while lower scores indicate less favorable concentrations. The right panel shows cumulative hazard curves for ARDs during follow-up, stratified by baseline life-course social standing scores. Curve separation by social standing scores, derived from cumulative logistic regression, became apparent after baseline and widened over the 20-year follow-up. The incidence of ARD observed in individuals with the highest social standing score at follow-up year 20 was reached 5.3 years earlier in those with the lowest score. **c**, Age-, sex- and ethnicity-adjusted  $\beta$  coefficients from cumulative logistic regression analysis for the 11 proteins available in the ARIC study confirmed the associations with social disadvantage observed in the Whitehall study when analyzing midlife protein concentrations. With two exceptions, these associations were also evident in the analysis of protein levels measured in old age. **d**, The forest plot shows age-, sex- and ethnicity-adjusted hazard ratios per 1 s.d. higher protein concentration for mortality, based on Cox proportional hazards regression analyses. The whiskers represent 95% CIs.

groups with different social standing scores emerged shortly after baseline and widened throughout the entire follow-up period (Fig. 4b). Notably, the 20-year incidence of ARDs observed in individuals with the highest social standing score was reached more than 5 years earlier among those with the lowest score, highlighting age-related acceleration in ARD incidence in the latter group.

**Replication of protein associations in an independent cohort**  
To examine whether the findings on plasma proteins were replicable in an independent study population, we repeated the main analyses using data from the ARIC study. Like the Whitehall study, the ARIC study includes data on indicators of social disadvantage—such as education, neighborhood deprivation and family income—alongside protein measurements using the SomaLogic platform, with long-term follow-up for all-cause mortality. In ARIC, the SomaScan 4.0 assay measured 11 of the

14 proteins, with data collected during midlife (mean age 57.1) for 11,798 participants and in older age (mean age 75.7) for 5,195 participants (Supplementary Tables 27 and 28). These proteins were S100A12, CCL16, CRP, F2, S100A9, FABP3, DNAJB9, FURIN, SUMF2, HSPA1A and HSPA1B. Results from education and adult SES in midlife (neighborhood deprivation) in ARIC replicated all associations with plasma proteins (Fig. 4c and Supplementary Table 29). At older ages, the findings were directionally consistent for these proteins, with the exception of SUMF2, which was not associated with social disadvantage in the ARIC cohort (Supplementary Table 30). In addition, the associations between proteins and mortality were directionally similar in the Whitehall study (follow-up of 22.4 years, s.d. 2.6) and the ARIC study, including the middle-age cohort with a mean follow-up of 22.7 years (s.d. 8.4) and the older cohort with a mean follow-up of 8.1 years (s.d. 2.5) (Fig. 4d).

## Test of genetic explanations

In addition to social causation and health-related selection (reversed causation), genetic predisposition may also drive the associations between social disadvantage, protein levels and hallmark-specific ARDs. To test this, we used an established polygenic risk score (PRS) for education<sup>21</sup>, available in the UK Biobank and Whitehall studies. Adjusting for this PRS, along with age, sex and ethnicity, had little impact on the associations of education and adult SES with hallmark-specific ARDs in these cohorts (Supplementary Tables 31 and 32). Similarly, the associations between education, adult SES and the 14 proteins in the Whitehall study remained robust following this adjustment (Supplementary Table 33). This was also the case for the associations between the proteins and hallmark-specific ARDs (Supplementary Table 34).

These findings indicate that genetic variation linked to social disadvantage has only a modest impact on the relationships between social disadvantage, age-related proteins and hallmark-specific ARDs.

## Discussion

This multicohort analysis of social disadvantage in relation to circulating proteins, ARDs and multimorbidity suggests that social disadvantage adversely affects a web of primary, compensatory and integrative hallmarks of aging, modifying multiple biological systems rather than specific tissues or organs. The over-twofold increased risk of type 2 diabetes, liver disease, heart failure, stroke, peripheral artery disease, gout, chronic obstructive pulmonary disease, lung cancer and mesothelioma in individuals with social disadvantage highlights the strong link between social disadvantage and hallmark-related morbidity. Furthermore, the 14 identified age-related circulating proteins mediated up to 39% of the social-disadvantage-related excess risk of hallmark-specific ARDs. These proteins were enriched in immune system responses to internal and external stimuli, with key downstream pathways via over-activation of NF- $\kappa$ B and IL-8. Changes in social disadvantage altered the levels of these proteins, suggesting that they may be modifiable.

Our study examines organ-specific proteomic aging in relation to social disadvantage, identifies over 1,000 plasma proteins potentially linked to the hallmarks of aging and investigates 83 hallmark-related diseases associated with social disadvantage and their protein mediators. The relationship between social disadvantage and hallmark-specific ARDs was robust, consistently observed in three independent cohort studies in the United Kingdom and Finland and replicated across indicators of both early- and later-life social disadvantage. Consistent findings from protein analyses in long-term follow-up studies in the United Kingdom and the United States further strengthen this evidence.

The validity of our study is supported by concordant findings from well-characterized, smaller-scale studies focused on a subset of aging hallmarks<sup>25–31</sup>. They suggest socioeconomic gradients in DNA methylation clocks<sup>7–9</sup>, associations between low education and TA<sup>27–30</sup> and a link between low social rank and EA in mouse livers<sup>31</sup>. The overactivation of the NF- $\kappa$ B–IL-8 pathway identified in this study is associated with ARDs across multiple organ systems, potentially linking social disadvantage to hallmarks of aging<sup>25,26,32</sup>. Consistent with our findings, ‘inflamm-aging’<sup>12</sup>—a term denoting the accumulation of non-resolving low-grade inflammation—has been proposed as an additional meta-cellular hallmark of aging, interconnected with other aging hallmarks such as GI, EA, LOP, MD and CS<sup>33</sup>.

We found a modest contribution of accelerated organ-specific aging to the excess risk of ARDs in social disadvantaged individuals. However, this finding should be interpreted cautiously, as organ-specific aging was assessed using plasma protein signatures based on fewer than 120 proteins highly expressed in the specific organ<sup>24</sup>. Organ aging may involve processes outside the studied organs or processes not captured by the measured organ-specific gaps. Due to these limitations, we may have overlooked some socially patterned proteins contributing to accelerated organ aging.

The upstream drivers of the observed associations between social disadvantage, age-related proteins and ARDs remain unknown but probably include factors such as unhealthy behaviors<sup>34,35</sup>; life stress<sup>11</sup>; environmental exposures, such as pollutants and toxins<sup>36</sup>; and depression<sup>37</sup>. Research indicates that individuals experiencing social disadvantage are more likely to engage in smoking, heavy alcohol consumption and drug abuse; consume an energy-dense nutrient-poor diet; show physical inactivity; experience disturbed sleep; neglect safety behaviors (for example, self-examination, participation in screening and vaccination); and fail to adhere to medical procedures (for example, antihypertensive medication use in patients with hypertension)<sup>38,39</sup>. Our protein findings could help identify vulnerable individuals who are most likely to benefit from interventions aimed at reducing social inequalities by targeting these factors.

While the findings regarding the 14 socially patterned, age-related proteins from the Whitehall study are promising, they require further validation in future research. We successfully replicated associations for 11 of these proteins in an independent cohort study, linking them to various social adversity indicators and mortality; however, data for the remaining 3 proteins were unavailable. Further research is also needed to identify disease profiles for the expanded hallmarks of aging, including impaired macroautophagy, chronic inflammation and dysbiosis, in addition to the original nine hallmarks investigated in our study<sup>33</sup>. In addition, our use of electronic health records shares limitations common to most large multi-outcome cohort studies. These include the inability to detect preclinical and undiagnosed conditions, as well as diseases that rarely result in hospitalization.

Our data originated from countries with distinct social contexts shaped by differences in welfare systems, economic inequality and service access<sup>40,41</sup>. Finland’s welfare state ensures universal healthcare, free education and strong social safety nets, promoting high income equality. The United Kingdom offers a moderately comprehensive system, including the National Health Service (NHS), but faces greater income inequality and entrenched class structures. The United States, with a less comprehensive welfare system, has a dynamic economy but significant disparities between affluent and poorer communities. While data collection methods varied across countries, limiting direct comparisons, the consistent findings across these diverse contexts suggest that our main conclusions may be generalizable beyond the studied populations.

In conclusion, our data support the concept that social disadvantage accelerates the aging process, leading to earlier age-related deterioration in affected individuals. This concept offers a potential explanation for the fundamental role that social stratification plays in determining our health. Our study suggests that these effects may be measurable through plasma and disease profiles.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-025-03563-4>.

## References

1. Stringhini, S. et al. Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. *Lancet* **389**, 1229–1237 (2017).
2. Snyder-Mackler, N. et al. Social determinants of health and survival in humans and other animals. *Science* **368**, eaax9553 (2020).
3. Kivimaki, M. et al. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. *Lancet Public Health* **5**, e140–e149 (2020).

4. Fraser, H. C. et al. Biological mechanisms of aging predict age-related disease co-occurrence in patients. *Aging Cell* **21**, e13524 (2022).
5. Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).
6. Kuan, V. et al. Data-driven identification of ageing-related diseases from electronic health records. *Sci. Rep.* **11**, 2938 (2021).
7. Schmitz, L. L. et al. The socioeconomic gradient in epigenetic ageing clocks: evidence from the Multi-Ethnic Study of Atherosclerosis and the Health and Retirement Study. *Epigenetics* **17**, 589–611 (2022).
8. Avila-Rieger, J. et al. Socioeconomic status, biological aging, and memory in a diverse national sample of older US men and women. *Neurology* **99**, e2114–e2124 (2022).
9. Fiorito, G. et al. Social adversity and epigenetic aging: a multi-cohort study on socioeconomic differences in peripheral blood DNA methylation. *Sci. Rep.* **7**, 16266 (2017).
10. Barzilay, N., Huffman, D. M., Muzumdar, R. H. & Bartke, A. The critical role of metabolic pathways in aging. *Diabetes* **61**, 1315–1322 (2012).
11. Kivimäki, M., Bartolomucci, A. & Kawachi, I. The multiple roles of life stress in metabolic disorders. *Nat. Rev. Endocrinol.* **19**, 10–27 (2022).
12. Ferrucci, L. & Fabbri, E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. *Nat. Rev. Cardiol.* **15**, 505–522 (2018).
13. Furman, D. et al. Chronic inflammation in the etiology of disease across the life span. *Nat. Med.* **25**, 1822–1832 (2019).
14. Alpert, A. et al. A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. *Nat. Med.* **25**, 487–495 (2019).
15. Berger, E. et al. Multi-cohort study identifies social determinants of systemic inflammation over the life course. *Nat. Commun.* **10**, 773 (2019).
16. Stringhini, S. et al. Association of lifecourse socioeconomic status with chronic inflammation and type 2 diabetes risk: the Whitehall II prospective cohort study. *PLoS Med.* **10**, e1001479 (2013).
17. Noppert, G. A., Stebbins, R. C., Dowd, J. B. & Aiello, A. E. Socioeconomic and race/ethnic differences in immunosenescence: evidence from the Health and Retirement Study. *Brain Behav. Immun.* **107**, 361–368 (2022).
18. Ferrucci, L. et al. Measuring biological aging in humans: a quest. *Aging Cell* **19**, e13080 (2020).
19. Rutledge, J., Oh, H. & Wyss-Coray, T. Measuring biological age using omics data. *Nat. Rev. Genet.* **23**, 715–727 (2022).
20. Lehallier, B., Shokhirev, M. N., Wyss-Coray, T. & Johnson, A. A. Data mining of human plasma proteins generates a multitude of highly predictive aging clocks that reflect different aspects of aging. *Aging Cell* **19**, e13256 (2020).
21. Okbay, A. et al. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat. Genet.* **54**, 437–449 (2022).
22. Kuan, V. et al. A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service. *Lancet Digit. Health* **1**, e63–e77 (2019).
23. Kivimäki, M. et al. Obesity and risk of diseases associated with hallmarks of cellular ageing: a multicohort study. *Lancet Healthy Longev.* **5**, e454–e463 (2024).
24. Oh, H. S. et al. Organ aging signatures in the plasma proteome track health and disease. *Nature* **624**, 164–172 (2023).
25. Liu, T., Zhang, L., Joo, D. & Sun, S. C. NF-κB signaling in inflammation. *Signal Transduct. Target. Ther.* **2**, 17023 (2017).
26. Ortiz-Montero, P., Londono-Vallejo, A. & Vernet, J. P. Senescence-associated IL-6 and IL-8 cytokines induce a self- and cross-reinforced senescence/inflammatory milieu strengthening tumorigenic capabilities in the MCF-7 breast cancer cell line. *Cell Commun. Signal.* **15**, 17 (2017).
27. Steptoe, A. et al. Educational attainment but not measures of current socioeconomic circumstances are associated with leukocyte telomere length in healthy older men and women. *Brain Behav. Immun.* **25**, 1292–1298 (2011).
28. Coimbra, B. M. et al. A systematic review on the effects of social discrimination on telomere length. *Psychoneuroendocrinology* **120**, 104766 (2020).
29. Willis, M., Reid, S. N., Calvo, E., Staudinger, U. M. & Factor-Litvak, P. A scoping systematic review of social stressors and various measures of telomere length across the life course. *Ageing Res. Rev.* **47**, 89–104 (2018).
30. Oliveira, B. S. et al. Systematic review of the association between chronic social stress and telomere length: a life course perspective. *Ageing Res. Rev.* **26**, 37–52 (2016).
31. Razzoli, M., Nyuyki-Dufe, K., Chen, B. H. & Bartolomucci, A. Contextual modifiers of healthspan, lifespan, and epigenome in mice under chronic social stress. *Proc. Natl Acad. Sci. USA* **120**, e2211755120 (2023).
32. Tilstra, J. S., Clauson, C. L., Niedernhofer, L. J. & Robbins, P. D. NF-κB in aging and disease. *Aging Dis.* **2**, 449–465 (2011).
33. Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. Hallmarks of aging: an expanding universe. *Cell* **186**, 243–278 (2023).
34. Raffin, J. et al. Sedentary behavior and the biological hallmarks of aging. *Ageing Res. Rev.* **83**, 101807 (2022).
35. Kankaanpää, A. et al. The role of adolescent lifestyle habits in biological aging: a prospective twin study. *eLife* **11**, e80729 (2022).
36. Racková, L., Mach, M. & Brnoliaková, Z. An update in toxicology of ageing. *Environ. Toxicol. Pharmacol.* **84**, 103611 (2021).
37. Lorenzo, E. C., Kuchel, G. A., Kuo, C. L., Moffitt, T. E. & Diniz, B. S. Major depression and the biological hallmarks of aging. *Ageing Res. Rev.* **83**, 101805 (2022).
38. Petrovic, D. et al. The contribution of health behaviors to socioeconomic inequalities in health: a systematic review. *Prev. Med.* **113**, 15–31 (2018).
39. Kraft, P. & Kraft, B. Explaining socioeconomic disparities in health behaviours: a review of biopsychological pathways involving stress and inflammation. *Neurosci. Biobehav. Rev.* **127**, 689–708 (2021).
40. OECD. *Income Inequalities* <https://www.oecd.org/en/data/indicators/income-inequality.html> (2021).
41. Lozano, R. et al. Measuring universal health coverage based on an index of effective coverage of health services in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1250–1284 (2020).

**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/>.

© The Author(s) 2025

## Methods

### Study design and participants

We analyzed pseudonymized individual-level data from four prospective cohort studies: the UK Biobank, FPS, Whitehall II and ARIC, with appropriate approvals for access.

The UK Biobank study is a nationwide, prospective cohort study of half of a million participants aged 38–73 years, living in the United Kingdom (<https://www.ukbiobank.ac.uk>)<sup>42</sup>. Baseline clinical examinations, including measures of social disadvantage, were conducted between 2006 and 2010. Participant follow-up via linked electronic health records was conducted until 2021. Our analyses were conducted using the UK Biobank Resource under application numbers 60565 and 22627. Data collection in the UK Biobank was approved by the North West Multi-Centre Research Ethics Committee. All participants provided written informed consent before their involvement in the study.

The FPS is a prospective occupational cohort study<sup>3</sup>. Study participants included 286,475 men and women aged 40 or older, who were followed up via electronic health records of national health registries until 2016. These registries provide dates and diagnoses of in-patient care in hospitalizations and deaths for all residents in the country. Participants provided electronic informed consent for both baseline assessments and register linkage. The FPS has received ethical approval from the ethical committee of the Helsinki and Uusimaa hospital district (HUS/1210/2016).

The Whitehall study is a prospective occupational cohort study of 10,308 London-based office staff, aged 35–55, with a baseline examination conducted between 1985 and 1988<sup>43</sup>. Blood samples for proteomic analyses were collected for 6,545 participants between 1997 and 1999, all of whom were linked to NHS electronic health records. Written informed consent from participants was obtained at each contact, and ethics approval was obtained from the University College London Hospital Committee on the Ethics of Human Research (reference number 85/0938).

The ARIC study is a population-based cohort study of 15,792 participants aged 45–64 years at baseline (1987–1989), selected through probability sampling from 4 US communities (Washington County, Maryland; Forsyth County, North Carolina; the Minneapolis, Minnesota, suburbs; Jackson, Mississippi)<sup>44</sup>. In proteomic analyses, self-identified non-Black and non-White participants and self-identified Black participants at Washington County and Minneapolis study sites were excluded owing to small sample sizes ( $n = 78$  at visit 2 and  $n = 29$  at visit 5). Blood samples for proteomic analyses were available for 11,798 participants during visit 2 (1990–1992) and for 5,195 participants during visit 5 (2011–2013). Mortality was ascertained through multiple data sources up until 31 December 2021. The study was approved by each field center's institutional review board, and all participants provided written informed consent before the study.

A flowchart of sample selection and baseline characteristics of the participants are presented in Extended Data Figs. 1 and 2, and Supplementary Table 1.

### Measurement of social disadvantage

Social disadvantage at baseline was assessed using several indicators: education, father's occupational status, neighborhood deprivation, occupational position and household income. In the UK Biobank and the Whitehall study, educational attainment was self-reported and classified into three categories: high (college or university degree), intermediate (A levels and AS levels or equivalent, O levels and GCSEs or equivalent, CSEs or equivalent, NVQ or HND or HNC or equivalent, other professional qualifications, for example, nursing, teaching), low (none of the above). In the FPS, information about education was obtained from Statistics Finland via record linkage, including the following three categories: high (tertiary qualification, college or university), intermediate (all other educational qualifications), low

(no qualifications, compulsory schooling). In ARIC, self-reported education was categorized as high (any college), intermediate (high school/GED/vocational school) and low (less than high school).

In the UK Biobank and the Whitehall study, neighborhood deprivation was assessed using the Townsend index<sup>45</sup> linked to the participants' residential address. This index is calculated from census data on four variables: unemployment (as a percentage of those aged 16 and over who are economically active), non-car ownership (as a percentage of all households), non-home ownership (as a percentage of all households) and household overcrowding. We divided the neighborhood deprivation score into three categories: high (top quartile), intermediate (middle quartiles) and low (bottom quartile).

For each FPS participant, the residential address was linked to a Statistics Finland score for neighborhood deprivation. The score is derived from three components: the proportion of adults with low education, the unemployment rate and the proportion of people living in rented housing within each 250 m by 250 m grid area<sup>3</sup>. As previously, we categorized these data as low (an area deprivation score lower than the national mean), intermediate (from the national mean to the national mean plus 0.5 s.d.) and high (higher than 0.5 s.d.)<sup>3</sup>.

In the ARIC study, neighborhood deprivation was measured using the National Level Area Deprivation Index, which was derived by linking geocoding participant's geocoded address coordinates with the US Census block group boundaries. Each census block was assigned a rank ranging from 1 to 100, calculated from 18 census-based indicators of socioeconomic disadvantage using the Singh method<sup>46</sup>. We divided the index into tertiles, to represent low, intermediate and high levels of neighborhood deprivation.

In the FPS, information about occupational position was coded based on the International Standard Classification of Occupations (ISCO)<sup>47</sup> and categorized into three occupational position groups: high (non-manual occupations, ISCO classes 1 and 2, for example, physicians, lawyers), intermediate (non-manual occupations, ISCO classes 3 and 4, for example, registered nurses) and low (service and manual occupations, ISCO classes 5–9, for example, cleaners, maintenance workers). In the Whitehall study, occupational position was obtained from the British civil service occupational grade at baseline, a 3-level variable representing high (administrative), intermediate (professional or executive) and low (clerical or support) grades<sup>43</sup>. In addition, father's occupational class was assessed with the question 'What is/was your father's main job, what kind of work does/did he do in it'. Responses were coded as manual or non-manual.

### PRS for education

The UK Biobank genetic data include genotypes for 488,377 participants, assayed using two very similar genotyping arrays: 807,411 markers using the Applied Biosystems UK BiLEVE Axiom Array by Affymetrix (now part of Thermo Fisher Scientific, for a subset of 49,950 participants) and 825,927 markers using the closely related Applied Biosystems UK Biobank Axiom Array (438,427 participants; shares 95% of marker content with the UK BiLEVE Axiom Array)<sup>48</sup>. In the Whitehall study, genotyping data include the Illumina Human Drug Core Array, which features a whole-genome single-nucleotide polymorphism scaffold with enhanced coverage of 200,000 custom markers in 4,500 drugged or druggable genes<sup>49</sup>.

We constructed a polygenic index for education, as defined in ref. 21, to examine whether the associations between social disadvantage, hallmark-related diseases and age-related proteins are attributable to genetic effects across all these constructs.

### Follow-up for ARDs and mortality

UK Biobank and Whitehall participants were linked to the UK NHS Hospital Episode Statistics database for hospital admissions and the NHS Central Registry for mortality. Electronic health records, including dates and the International Classification of Diseases diagnostic

codes of hospitalizations and deaths, were retrieved until 2021 in the UK Biobank study and until 2019 in the Whitehall study. The NHS provides most of the healthcare in the country, including in- and out-patient care, and record linkage was undertaken using a unique NHS identifier held by all UK residents. FPS participants were linked by their unique identification number to national hospital discharge (recorded by the Finnish Institute for Health and Welfare) and mortality (recorded by Statistics Finland) registries. These electronic health records included cause (International Classification of Diseases codes) and date of hospitalization or mortality, or both, until 2016.

In the UK Biobank, FPS and Whitehall study, we measured the 4 primary and 5 compensatory and integrative hallmarks of aging indirectly, based on a person's vulnerability to specific hallmark-related diseases as defined in refs. 4,22. This list comprises a total of 85 diseases, including those 30 diseases most strongly related to each hallmark. These diseases encompass conditions linked to a single hallmark and also those shared by two or more hallmarks. The presence of a specific hallmark of aging is indicated if the participant has one or more diseases related to that hallmark. The 83 diseases and their diagnostic codes for each aging hallmark that were available for this study are listed in Supplementary Table 1 (ref. 23).

An extended model of aging hallmarks includes 3 additional hallmarks: disabled macroautophagy (a special case of LOP in the original model), chronic inflammation (a component of AIC in the original model) and dysbiosis (also part of AIC in the original model)<sup>33</sup>. A pre-defined validated list of diseases related to these 3 new hallmarks was not available at the time of our study.

In the ARIC study, mortality was ascertained through contact with participant proxies via telephone, hospital records, death certificates or vital statistics from the National Death Index until 31 December 2021. As unified national hospitalization registries with comprehensive coverage, such as those in the United Kingdom and Finland, are not available in the United States, similarly high-quality data on hallmark-related diseases were not available in the ARIC study.

### Statistical power and reproducibility

Sample sizes were determined based on available data in the cohorts. As shown in Supplementary Figs. 1 and 2, participants were excluded from the analyses owing to missing data on social disadvantage or hallmark-related diseases. A total of 492,257 participants were included in the analysis of the UK Biobank and 286,475 participants in the analysis of the FPS. In prospective analyses, participants with hallmark-related diseases at or before baseline were additionally excluded. The number of included participants varied between 430,307 and 452,305 in the UK Biobank and between 267,046 and 281,348 in the FPS, depending on the hallmark under investigation (Supplementary Table 5).

Statistical power varied depending on the outcome. For any incident hallmark-related disease in the UK Biobank, we were able to detect a 3–4% difference between high- and low-education groups (hazard ratios 1.03–1.04) at 90% power and an alpha level of  $P = 0.05$ . The corresponding range for hazard ratios was 6–7% (hazard ratios 1.06–1.07) in the FPS. For specific hallmark-related diseases, the minimally detectable difference in hazard ratios ranged from 1.05 (cataract and osteoarthritis) to 2.65 (essential tremor and scleroderma) in the UK Biobank and from 1.10 (cataract and osteoarthritis) to 3.75 (essential tremor and Barrett's esophagus) in the FPS.

Reproducibility was assessed using multiple indicators of social disadvantage across four independent cohort studies from the United Kingdom, Finland and the United States. Further examination of reproducibility used methodological triangulation, that is, examining the consistency of the results across alternative epidemiological approaches and sensitivity analyses. The main finding of the social disadvantage–aging hallmark association was consistent across the three independent cohort studies, four alternative indicators of social disadvantage and in multiple alternative analyses.

### Measurement of plasma proteins

Plasma ethylenediaminetetraacetic acid (EDTA) samples drawn at baseline, coinciding with the assessment of social disadvantage, were stored in 0.25-ml aliquots at  $-80^{\circ}\text{C}$ . Proteins were analyzed using the SomaScan version 4.0 and 4.1 assays by SomaLogic, including up to over 7,000 unique proteins. We used the SOMAmer-based capture array, which quantifies the relative concentration of plasma proteins or protein complexes. The SomaScan platform uses short single-stranded DNA with chemically modified nucleotides (modified aptamers) that act as protein-binding reagents with defined three-dimensional structures and unique nucleotide sequences. These aptamers are identifiable and quantifiable using DNA detection technology. This measurement was blinded to participant characteristics, including social disadvantage or hallmark-related disease.

### SomaLogic normalization and quality control

All samples underwent standard SomaLogic normalization, calibration and quality control. To control for batch effects during assay quantification, pooled reference standards and buffer standards are included on each plate. To control for both within-plate and across-plate technical variation, samples are normalized within and across plates using median signal intensities in reference standards. Samples are further normalized to a pooled reference using an adaptive maximum likelihood procedure. If signal intensities deviate significantly from the expected range, samples are additionally flagged by SomaLogic and these samples were excluded from the analysis. SomaLogic provides data on the resulting expression values (the 'raw' data).

The scale factor acceptance criterion per plate is between 0.4 and 2.5, and this criterion was passed for all 48 plates for Whitehall EDTA samples. The 'calibrator percent-in-tails' refers to the percentage of plate calibration scale factors with values outside the expected range, 0.6–1.4. The alert criterion of 10% was slightly exceeded in 2 of the 48 plates (10.8% and 10.9%, respectively), the values for other plates ranging between 1.7% and 9.6%. The 'QC percent-in-tails' refers to the percentage of SOMAmer reagents in the QC control that are outside the accepted accuracy range, 0.8–1.2, when compared with the reference. All plates in the Whitehall study passed the acceptance criteria.

We applied the version 4.0 → version 4.1 multiplication scaling factors provided by SomaLogic to the raw version 4.0 assay expression values to allow direct comparisons across samples analyzed using version 4.0 and version 4.1. The raw data were transformed to a normal distribution by inverse rank-based normal transformation before analysis, as the assay has an expected log-normal distribution.

### SomaLogic probe validation

SomaLogic SomaScan assay technology has been widely used in biomedical research, the list of publications comprising ~700 peer-reviewed papers (<https://somalogic.com/publications/>). The performance of the SomaScan assay and the modified aptamer binding has been described in detail elsewhere<sup>50,51</sup>. Briefly, there is minimal replicate sample variability (coefficient of variation). The specificity of aptamer reagents is good and has been confirmed in several ways. The median intra- and inter-assay coefficients of variation for SomaScan are ~5%, and assay sensitivity is comparable to that of typical immunoassays, with a median lower limit of detection in the femtomolar range. The majority of SomaScan protein measurements are stable, and a subset of proteins has been validated using laboratory-developed tests. These validated proteins have been delivered from SomaLogic's Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory to physicians and patients in the context of medical management.

All 7,524 probes on the version 4.1 assay undergo rigorous primary validation of binding and sensitivity to the target protein, including determination of the equilibrium binding affinity dissociation constant, pull-down assay of cognate protein from buffer, demonstration of dose response in the SomaScan assay and estimation of endogenous

cognate protein signals in human plasma above the limit of detection. A total of 70% of the SomaScan probes have at least one orthogonal source of validation from mass spectrometry, antibody-based measurements, *cis*-protein quantitative trait loci analysis, absence of binding with the nearest neighbor (that is, no detected signal from the protein that is most closely related in sequence to the cognate protein) or correlation with mRNA levels in cell lines.

### Measurement of organ-specific aging signatures

We computed a total of 9 organ-age gaps, accounting for cohort characteristics, using the 'organage package'<sup>24</sup> in Python (<https://github.com/hamiltonoh/organage>). This package requires SomaScan data version 4.0 or version 4.1, and age and sex as inputs, to compute z-scores (mean = 0, s.d. = 1) for organismal and organ-specific age gaps, that is, the biological age of an individual's organs or body relative to those of same-aged peers. The resulting variables relate to the organ-specific ages of the arteries (14 proteins), brain (202 proteins), heart (10 proteins), immune function (173), intestine (33 proteins), kidneys (12 proteins), liver (113 proteins), lungs (9 proteins) and pancreas (34 proteins), in addition to an overall organismal age (3,907 proteins). The list of proteins included in each age gap is provided in Supplementary Table 9.

### Identification of proteins related to hallmarks of aging

We identified plasma proteins related to each hallmark of aging using the Human Proteomic Atlas ([www.proteinatlas.org](http://www.proteinatlas.org)) with the expanded 65-term taxonomy in ref. 4:

Genomic instability, mtDNA mutations, mtDNA damage, transposable elements, DNA damage, DNA repair deficiencies, mutations, DNA breaks, ssDNA breaks, dsDNA breaks, chromosome breakage; Telomere attrition, decreased telomere length, decreased leukocyte telomere length;  
Epigenetic alterations, gene transcription, coding-RNA, noncoding RNA, microRNA, DNA methylation, histone modifications, histone acetylation, histone methylation;  
Loss of proteostasis, endoplasmic reticulum stress, unfolded protein response, proteolysis, proteasome, autophagy, protein aggregation, chaperone;  
Deregulated nutrient sensing, insulin resistance, dyslipidaemia, nutrient sensing pathways, sirtuin 1, Insulin/insulin-like growth factor-1 signaling, mTORC1, AMP-activated protein kinase;  
Mitochondrial dysfunction, mitochondrial toxicity, reactive oxygen species, mitochondrial bioenergetics, electron transport chain, Krebs cycle, mitochondrial dynamics, mitochondrial turnover, mitochondrial degradation, mitochondrial biogenesis;  
Cellular senescence, senescence markers, senescence-associated secretory phenotype (SASP), immune-senescence;  
Stem cell exhaustion, stem cell differentiation, progenitor cell, stem cell self-renewal;  
Altered intercellular communication, inflammatory signaling, inflammaging, inflammation, neural signaling, neurotransmitters, hormonal signaling, hormones.

Details of the search are provided in Supplementary Table 12. To confirm the link with age and controlling for multiple testing, we included only those proteins that were additionally associated with chronological age after Bonferroni adjustment ( $P < 6.58 \times 10^{-6}$ ), a total of 1,044 proteins (Supplementary Table 13). Of these, 43 were related to SCE, 61 to AIC, 253 to DNS, 11 to CS, 159 to MD, 274 to EA, 1 to TA, 438 to LOP and 296 to GI (Supplementary Table 14).

### Reproducibility of proteomic findings

To confirm the reproducibility of our findings on protein–social disadvantage associations, we repeated the discovery analysis using low occupational position as the indicator of social disadvantage instead

of low education; only findings that were replicated across the different indicators of social disadvantage are reported. If multiple aptamers were available to determine a protein hit, we confirmed that the findings were consistent across all the aptamers.

To assess the reproducibility of the associations between proteins and social disadvantage, as well as between proteins and mortality, findings from the UK Whitehall study were validated using an independent dataset from the United States, the ARIC cohort study. Extending the Whitehall findings, the main analyses were repeated separately for individuals with protein measurements taken during midlife and those with measurements taken during old age.

### Statistical approach

In the examination of hallmark-related morbidities, data from the UK Biobank and the FPS were analyzed separately. In a retrospective analysis, we computed the rate of hallmark-specific diseases per 100 person-years for each aging hallmark and examined their rate ratios and 95% CIs by each index of social disadvantage using Poisson regression adjusted for age, sex and ethnicity. We modeled these rates and rate ratios for age 70 in the UK Biobank and for age 55 in the FPS. These were the mean ages at the end of follow-up in these studies.

In prospective analyses of participants with no hallmark-related diseases at baseline, we used Cox proportional hazards models to examine the age-, sex- and ethnicity-adjusted associations of social disadvantage indicators with the first onset of hallmark-related disease within each aging hallmark. To control for genetic confounding, we included PRS for education as an additional covariate in the model.

We used the same approach to examine the age-, sex- and ethnicity-adjusted associations of social disadvantage indicators with the onset of the second ARD among individuals who had already developed one ARD. In addition, we examined the corresponding associations with the onset of the third ARD among those who had already two ARDs.

We calculated clustering coefficients (range 0–1; a higher coefficient indicates stronger connections between the diseases) for groups of diseases that related to a specific hallmark of aging by using the Barrat method of global network transitivity<sup>52</sup>.

We computed phi coefficients to produce a correlation matrix across nine dichotomous aging hallmark variables. For each hallmark, this variable was defined as having at least one versus none of the hallmark-related diseases after baseline.

To examine whether the associations between social disadvantage indicators and the first onset of hallmark-related disease within each hallmark were driven by specific hallmark-related diseases, we analyzed the associations of social disadvantage indicators with all 83 hallmark-related diseases in separate Cox models adjusted for age, sex and ethnicity. To identify the strongest associations, we computed the mean hazard ratio across the two studies and all indicators of social disadvantage weighted by the number of disease cases.

We computed a total of nine organ age gaps, accounting for cohort characteristics, using the 'organage package' in Python (<https://github.com/hamiltonoh/organage>)<sup>24</sup>. This package requires SomaScan data version 4.0 or version 4.1, and age and sex as inputs, to compute z-scores (mean = 0, s.d. = 1) for organismal and organ-specific age gaps, that is, the biological age of an individual's organs or body relative to those of same-age peers.

Before protein discovery analyses were conducted in the Whitehall study, proteins were transformed to a normal distribution using inverse rank-based normal transformation. We included those hallmark-related proteins that were associated with chronological age at proteome-wide significance (0.05/30,000,  $P = 1.67 \times 10^{-6}$ ) after adjustment for sex and ethnicity. We then examined associations between hallmark-related proteins and indicators of social disadvantage (low occupational status and low education) using logistic regression analysis adjusted for chronological age, sex and ethnicity (White versus non-White)

and corrected for multiple testing using the Bonferroni method ( $P = 0.05/1,044, P = 4.78 \times 10^{-5}$ ).

We examined the associations between social disadvantage and the first onset of hallmark-related diseases in the Whitehall study, as was done in the UK Biobank and FPS, including testing for genetic confounding. To investigate the association of change in social disadvantage between early and later adulthood with the levels of proteins, we compared individuals with low education and intermediate or high occupational status (upward social trajectory) with those with low education and low occupational status (persistently low trajectory), as well as individuals with high education and intermediate or low occupational status (downward trajectory) with those with high education and high occupational status (persistently high trajectory). Accumulation of risk was examined using age-, sex- and ethnicity-adjusted Cox models, treating ‘adult life course social standing’ as a five-category exposure and hallmark-related diseases as the outcome.

To confirm the association between hallmark-related proteins and total mortality in both Whitehall and ARIC, we performed a Cox proportional hazards analysis for proteins associated with age and social disadvantage, testing whether these proteins were also linked to mortality after adjustment for age, sex and ethnicity. We report only hallmark-related proteins that were consistently associated with chronological age, mortality and social disadvantage during early (low education) and later life (low occupational status).

To identify protein mediators of the associations between social disadvantage and hallmark-related diseases, we estimated the proportion of this association mediated by social disadvantage-related proteins in a subgroup of people with no hallmark-related diseases at baseline using Cox regression models. Specifically, we used an inverse odds ratio-weighted method to estimate the extent to which the exposure (here social disadvantage) and the mediator (proteins) act as if they are independent of each other, that is, how the exposure directly affects the outcome when excluding the mediator pathway<sup>53</sup>. The inverse odds ratio-weighted method allows simultaneous assessment of multiple mediators, as well as individual mediators, by regressing the set of mediators on the exposure. This approach is particularly well suited for causal mediation analysis in complex models such as ours.

Analyses were performed using R and SAS (9.4), Stata (17.0) and RStudio (2023.03.1). To identify biological processes enriched by the identified 14 proteins, we used the clusterProfiler package and Gene Ontology (GO)-term enrichment analysis in R (ref. 54). This approach uses a hypergeometric test and false discovery rate correction for enrichment analyses. To identify a protein interaction network for the 14 proteins, we used the STRINGdb package in R (ref. 55). Interactions were searched across the entire String database with a minimum required interaction score of 0.4. Graphs were generated in Excel and the BioRender platform (<https://www.biorender.com>).

### Reporting summary

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

### Data availability

Researchers registered with the UK Biobank can apply for access to the individual-level data by completing an application. This must include a summary of the research plan, data fields required, any new data or variables that will be generated and payment to cover the incremental costs of servicing an application (<https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>). In the FPS, pseudonymized individual-level questionnaire data can be shared by request to the investigators (jaana.e.pentti@helsinki.fi). Access to linked health records requires separate permission from Findata, the Health and Social Data Permit Authority in Finland (<https://findata.fi/en/permits/>). Pseudonymized individual-level data from the Whitehall study are available for sharing within the

scientific community. Bona fide researchers interested in accessing the pseudonymized data can apply through Dementias Platform UK (<https://www.dementiasplatform.uk>) or the Whitehall Scientific Committee (<https://www.ucl.ac.uk/epidemiology-health-care/research/epidemiology-and-publichealth/research/whitehall-ii/data-sharing>). ARIC proteomic data are available through the NHLBI Biologic Specimen and Data Repository Information Coordinating Center (<https://biolincc.nhlbi.nih.gov/studies/aric>). Additional requests for clinical or proteomic data from individual investigators may be submitted to the ARIC steering committees and will be reviewed to ensure that data can be shared without compromising patient confidentiality or breaching intellectual property restrictions. Participant-level demographic, clinical and proteomic data may be partially restricted based on previously obtained participant consent. Data-sharing restrictions may also be applied to ensure consistency with confidentiality or privacy laws and considerations (<https://sites.cscs.unc.edu/aric>). Source data are provided with this paper.

### Code availability

All code developed and used throughout this study has been made open source and is available via GitHub at <https://github.com/rmjdmj1/SES-ageing.git> and in Supplementary Information.

### References

- 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).
- Marmot, M. & Brunner, E. Cohort profile: the Whitehall II study. *Int. J. Epidemiol.* **34**, 251–256 (2005).
- The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am. J. Epidemiol.* **129**, 687–702 (1989).
- Townsend, P., Phillimore, P. & Beattie, A. *Health and Deprivation: Inequality and the North* (Routledge, 1988).
- Singh, G. K. Area deprivation and widening inequalities in US mortality, 1969–1998. *Am. J. Public Health* **93**, 1137–1143 (2003).
- International Labour Organization. *International Standard Classification of Occupations (ISCO)* (ILO Bureau of Statistics, 2008).
- Wain, L. V. et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir. Med.* **3**, 769–781 (2015).
- Finan, C. et al. The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* **29**, eaag1166 (2017).
- Tin, A. et al. Reproducibility and variability of protein analytes measured using a multiplexed modified aptamer assay. *J. Appl. Lab. Med.* **4**, 30–39 (2019).
- Candia, J., Daya, G. N., Tanaka, T., Ferrucci, L. & Walker, K. A. Assessment of variability in the plasma 7k SomaScan proteomics assay. *Sci. Rep.* **12**, 17147 (2022).
- Barrat, A., Barthelemy, M., Pastor-Satorras, R. & Vespignani, A. The architecture of complex weighted networks. *Proc. Natl Acad. Sci. USA* **101**, 3747–3752 (2004).
- Nguyen, Q. C., Osypuk, T. L., Schmidt, N. M., Glymour, M. M. & Tchetgen Tchetgen, E. J. Practical guidance for conducting mediation analysis with multiple mediators using inverse odds ratio weighting. *Am. J. Epidemiol.* **181**, 349–356 (2015).
- Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
- Franceschini, A. et al. STRING v9.1: protein–protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**, D808–D815 (2013).

## Acknowledgements

This research was supported by the Wellcome Trust Collaborative Award in Science and the Intramural Research Program of the National Institute on Aging (NIA). M.K. is supported by the Wellcome Trust (221854/Z/20/Z), the UK Medical Research Council (S011676, Y014154), the US NIA (R01AG056477) and the Research Council of Finland (350426). J.P. and S.T.N. are supported by the Research Council of Finland (350426). P.F. is supported by the Wellcome Trust (221854/Z/20/Z). F.L. is supported by the Intramural Research Program of the NIA. K.A.W. is supported by the Intramural Research Program of the NIA. J.V.L. is supported by the Research Council of Finland (339568) and the Päivikki and Sakari Sohlberg Foundation. Analysis of 2,240 plasma samples in the Whitehall study was funded by SomaLogic, Inc, as part of an earlier academic–industry partnership project<sup>54</sup>. The ARIC is carried out as a collaborative study supported by National Heart, Lung and Blood Institute (NHLBI) contracts (grant numbers 75N92022D00001, 75N92022D00002, 75N92022D00003, 75N92022D00004 and 75N92022D00005). The funders of the study had no role in study design, data collection, data analysis, data interpretation or the writing of the report.

## Author contributions

M.K., J.P., L.P. and J.V.L. were responsible for designing the study. M.K., with P.F. and J.V.L., wrote the first draft. All authors were generally involved in the program and commented on the paper. Specific contributions were as follows: J.P., J.V.L. and F.L. analyzed the data. J.V.L. identified SomaLogic proteins associated with hallmarks of aging. M.K., J.P. and S.T.N. were accountable for the data from the FPS. M.K. and A.S.-M. were accountable for the data from the Whitehall II study. F.L. and K.A.W. were accountable for the data from the ARIC

study. T.W.-C. advised on plasma proteomics and the signatures of organ-specific age gaps. L.P. advised on analyses, hallmarks of aging and age-related mechanisms and was involved in the interpretation of the findings.

## Competing interests

T.W.-C. has filed a patent application related to proteomic organ aging, is a co-founder and scientific advisor of Teal Omics Inc. and has received equity stakes. T.W.-C. is also a co-founder and scientific advisor of Alkagest Inc. and Qinotto Inc. and has received equity stakes in these companies. The other authors declare no competing interests.

## Additional information

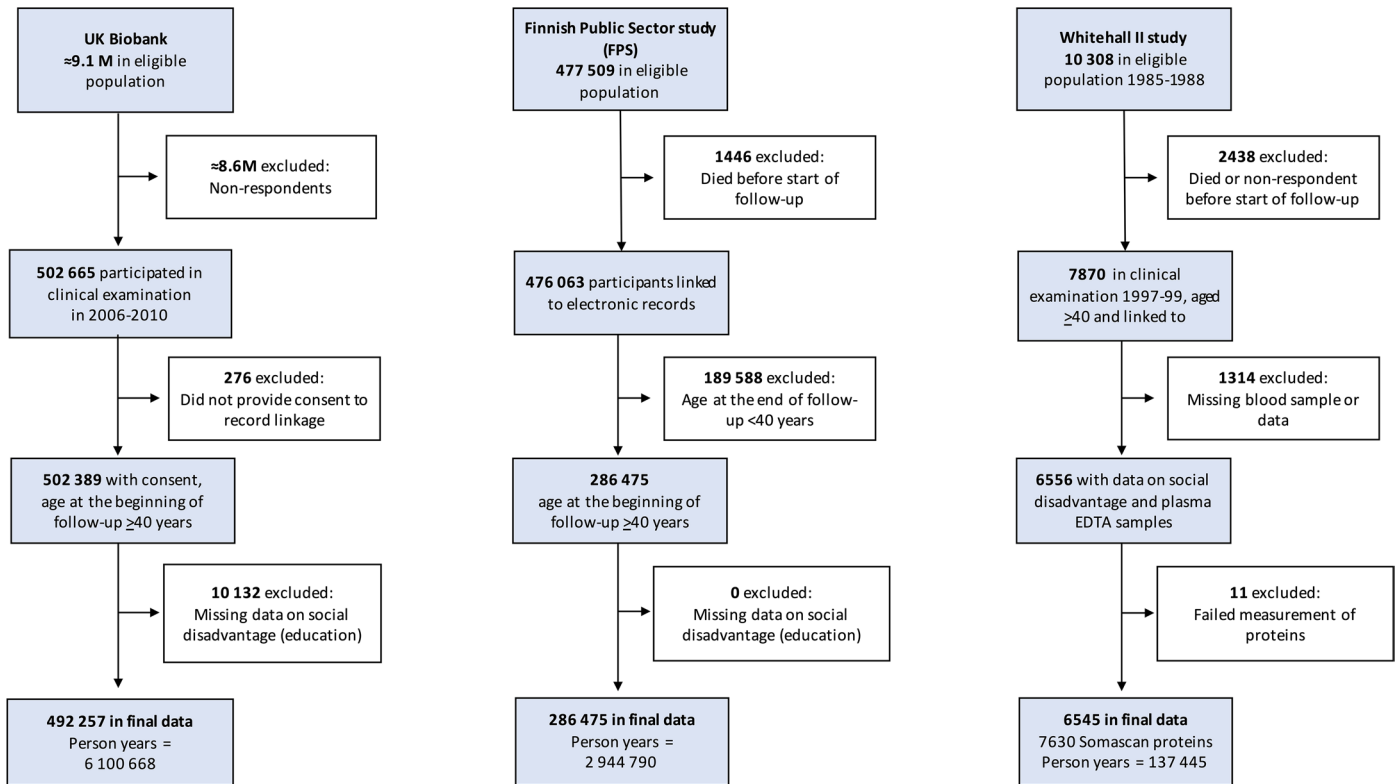
**Extended data** is available for this paper at <https://doi.org/10.1038/s41591-025-03563-4>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41591-025-03563-4>.

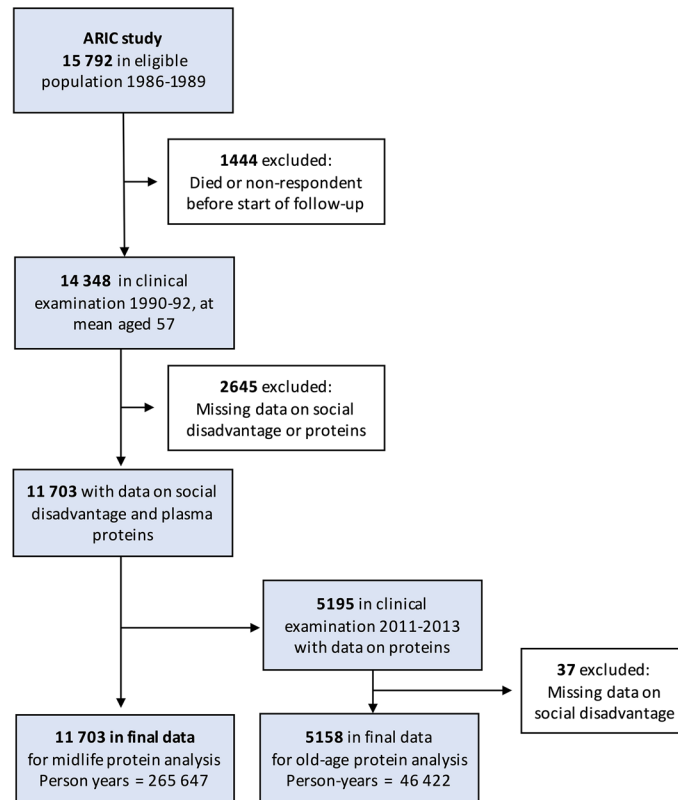
**Correspondence and requests for materials** should be addressed to Mika Kivimäki.

**Peer review information** *Nature Medicine* thanks Katsuya Oi, Cecilia Potente and Monica Williams-Farrelly for their contribution to the peer review of this work. Primary Handling Editor: Liam Messin, in collaboration with the *Nature Medicine* team.

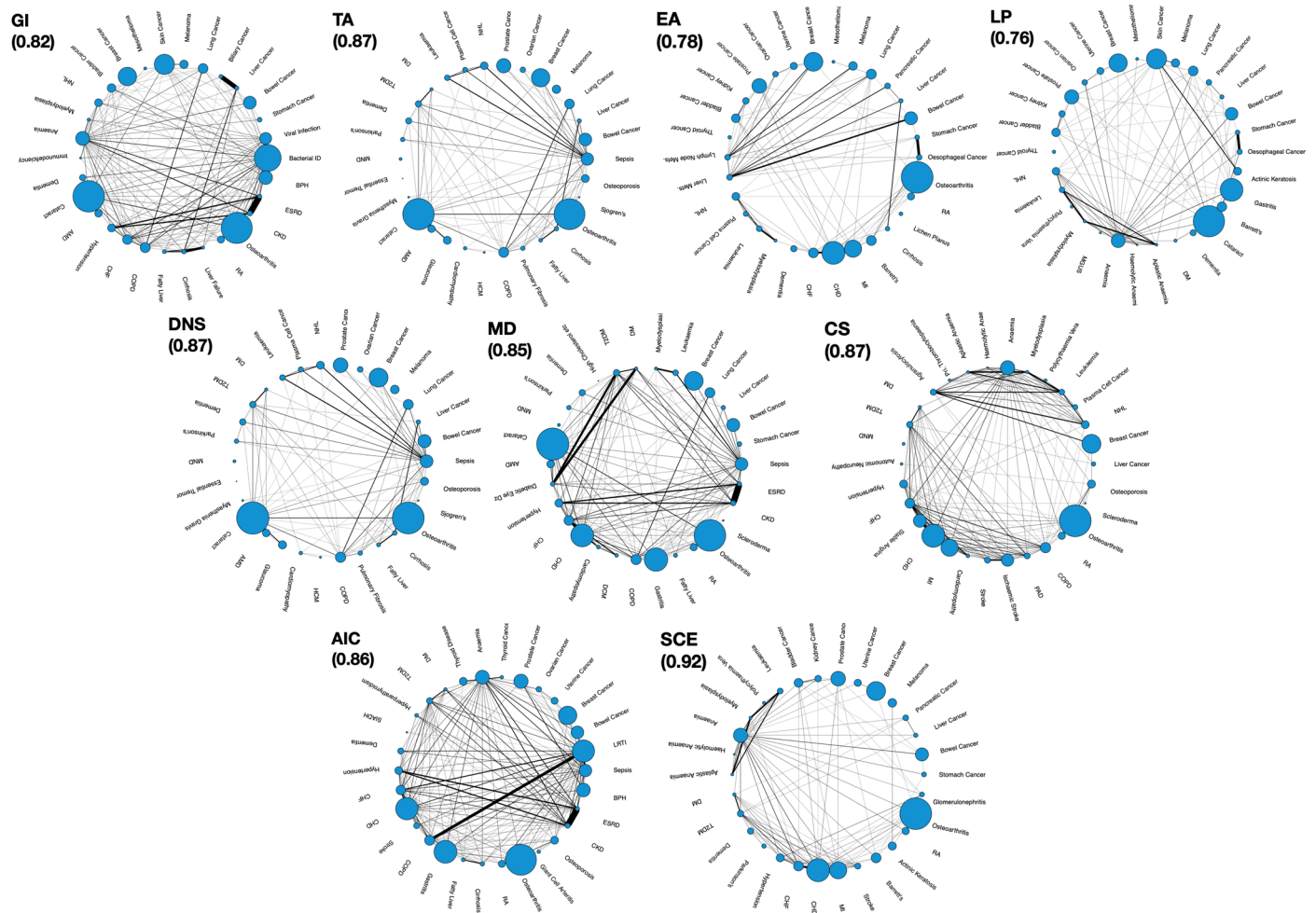
**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).



**Extended Data Fig. 1 | Fig. 1. Sample selection for the cohorts in the primary analysis.**



Extended Data Fig. 2 | Fig. 2 Sample selection for the cohorts in the replication study.



**Extended Data Fig. 3 | Fig. 3. Barrat Global Network Transitivity Analysis of disease associations in hallmark-specific ARDs (source: Kivimaki et al. *Lancet Healthy Longevity* 2024<sup>23</sup>). Data are from the UK Biobank study (total N = 492,257). Number in parenthesis is clustering coefficient. Circle size is proportional to the occurrence of the disease. AIC, altered intercellular**

communication; ARD, age-related disease; CS, cellular senescence; DNS, deregulated nutrient sensing; EA, epigenetic alterations; GI, genomic instability; LOP, loss of proteostasis; MD, mitochondrial dysfunction; SCE, stem cell exhaustion; TA, telomere attrition.

Hallmark-specific ARD	1	2	3	4	5	6	7	8	9
1. SCE	1.00	0.84	0.65	0.59	0.47	0.70	0.75	0.51	0.22
2. DNS	0.86	1.00	0.65	0.58	0.43	0.69	0.71	0.48	0.19
3. AIC	0.72	0.77	1.00	0.68	0.49	0.67	0.70	0.54	0.40
4. MD	0.74	0.71	0.73	1.00	0.66	0.59	0.58	0.75	0.55
5. CS	0.62	0.61	0.68	0.81	1.00	0.50	0.50	0.71	0.58
6. EA	0.78	0.78	0.76	0.72	0.65	1.00	0.89	0.60	0.41
7. TA	0.84	0.82	0.78	0.76	0.65	0.90	1.00	0.55	0.40
8. LOP	0.65	0.63	0.67	0.88	0.80	0.72	0.69	1.00	0.57
9. GI	0.35	0.36	0.42	0.62	0.61	0.52	0.48	0.69	1.00

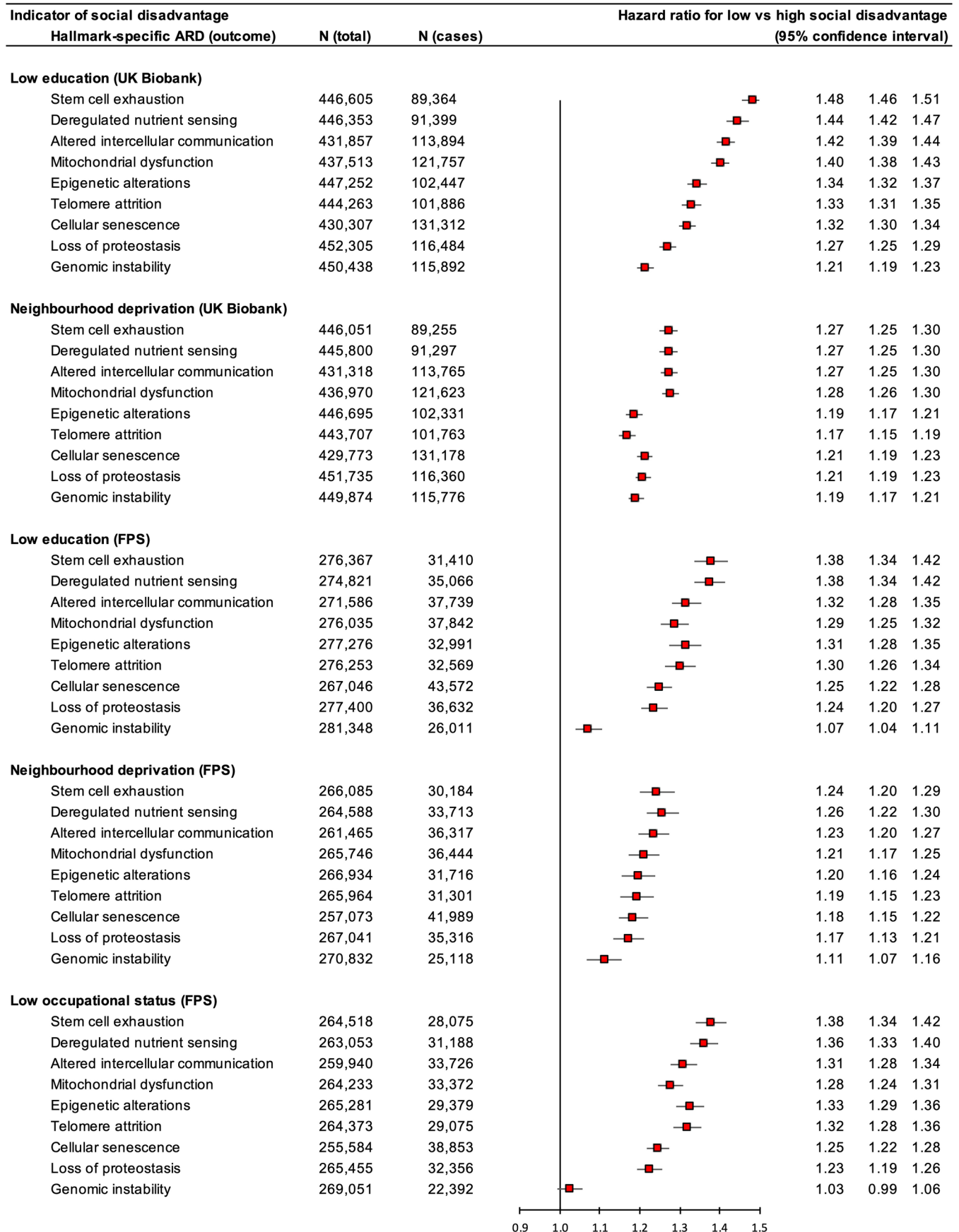
**Extended Data Fig. 4 | Fig. 4. Correlation matrix for hallmark-specific ARDs in the UK Biobank and the Finnish Public Sector studies.** Numbers are Phi coefficients of bivariate correlations between hallmark-specific ARDs in the UK Biobank (total N = 492,257) and Finnish Public Sector (total N = 286,475) studies.

AIC, altered intercellular communication; ARD, age-related disease; CS, cellular senescence; DNS, deregulated nutrient sensing; EA, epigenetic ageing; FPS, Finnish Public Sector study; GI, genomic instability; LOP, loss of proteostasis; MD, mitochondrial dysfunction; SCE, stem cell exhaustion; TA, telomere attrition.

Hallmark-related disease outcome	HR for high vs low social disadvantage				
	SCE	DNS	AIC	MD	CS
<b>Compensatory/integrative hallmarks</b>					
Stable angina	1.2-2.0				
Gout		1.7-2.8			
Lower respiratory track infection			1.5-2.3		
Diabetic eye disease				1.4-2.3	
Dilated cardiomyopathy				1.3-1.8	
Liver failure					2.0-3.5
Bacterial infection					1.2-1.6
<b>Primary hallmarks</b>	<b>EA</b>	<b>TA</b>	<b>LOP</b>	<b>GI</b>	
Glomerulonephritis		1.4-3.1			
Hypertension		1.4-1.7			
Fatty liver			1.7-3.1		
Sepsis			1.4-1.6		
Cardiomyopathy			1.3-2.0		
Gastritis				1.6-2.8	

**Extended Data Fig. 5 | Fig. 5. Cox proportional hazards regression analysis of social disadvantage and risk of single hallmark-specific ARDs.** Numbers represent the range of statistically significant age-, sex-, and ethnicity-adjusted hazard ratios for high versus low social disadvantage (education, neighbourhood deprivation, occupational position) in relation to specific ARDs in the UK Biobank (N = 492,257) and the Finnish Public Sector study (N = 286,475). These ranges reflect associations where ARDs are linked exclusively to a single compensatory/

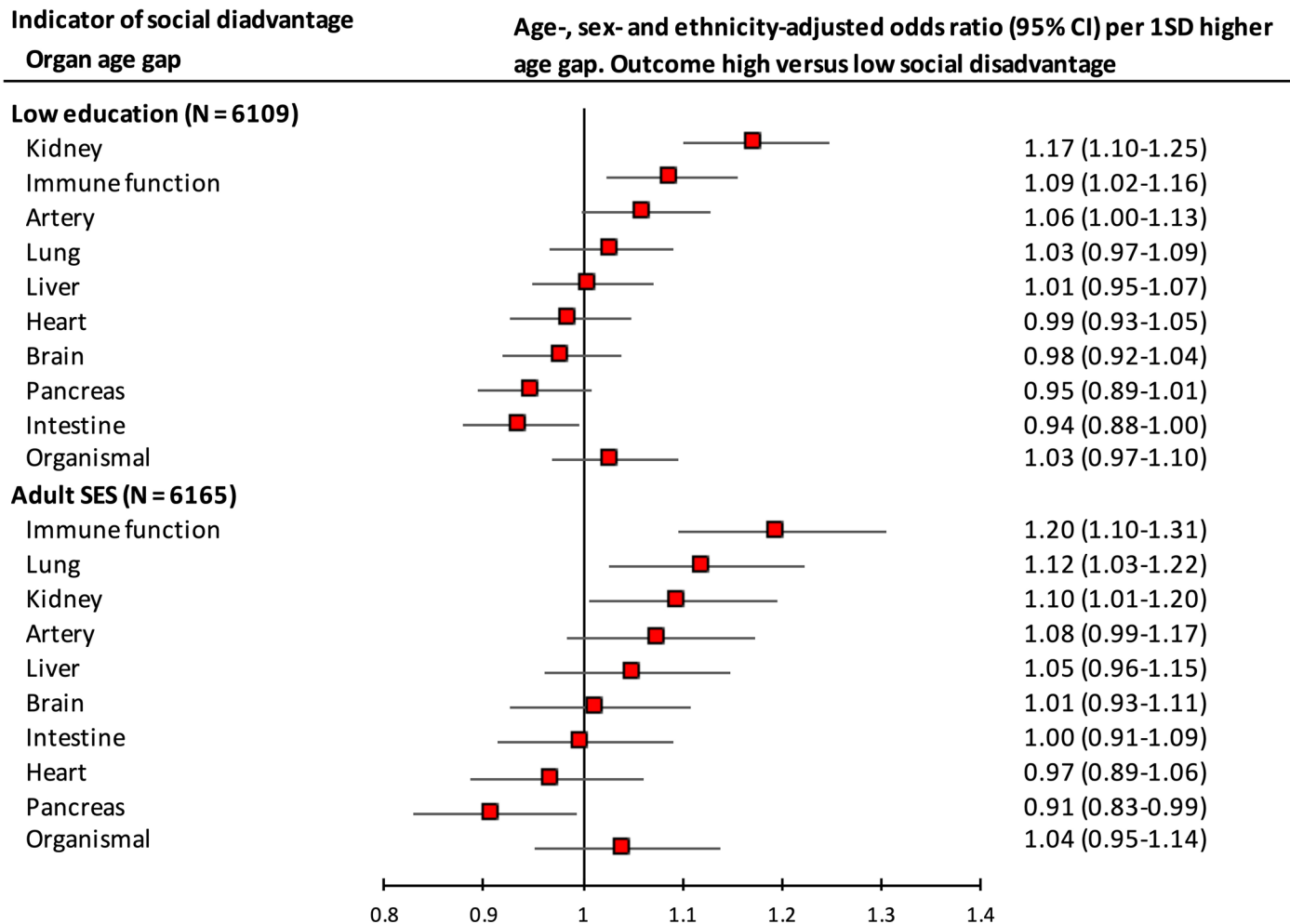
integrative hallmark or a single primary hallmark of aging. Full results for all social disadvantage-ARD associations are available in Supplementary Table 3. AIC, altered intercellular communication; ARD, age-related disease; CS, cellular senescence; DNS, deregulated nutrient sensing; EA, epigenetic ageing; GI, genomic instability; HR, hazard ratio; LOP, loss of proteostasis; MD, mitochondrial dysfunction; SCE, stem cell exhaustion; TA, telomere attrition.



Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 6 | Fig 6. Time-to-event analysis for social disadvantage and developing the first hallmark-specific ARDs in participants with no such diseases at the time social disadvantage was measured.** Data are from the UK Biobank and the Finnish Public Sector study. The plots show the age-, sex- and ethnicity-adjusted hazard ratios for hallmark-specific ARDs in participants with high versus low social disadvantage, based on Cox proportional hazards

regression analysis. Whiskers show 95% confidence interval for these estimates. AIC, altered intercellular communication; ARD, age-related disease; CS, cellular senescence; DNS, deregulated nutrient sensing; EA, epigenetic ageing; FPS, Finnish Public Sector study; GI, genomic instability; HR, hazard ratio; LOP, loss of proteostasis; MD, mitochondrial dysfunction; SCE, stem cell exhaustion; TA, telomere attrition.



**Extended Data Fig. 7 | Fig. 7. Cross-sectional associations between indicators of social disadvantage and plasma proteomic signatures of organ-specific and organismal (non-organ-specific) age gaps.** Data are from the Whitehall study. The plots display the odds ratios for low vs high social disadvantage from multinomial logistic regression analyses for organ-specific and organismal age gaps, defined as the biological age of an individual's organs or body relative to that of same-aged peers. Whiskers show 95% confidence interval for these estimates.

	S100A12	GCHFR	CCL16	CRP	CFHR1	F2	S100A9	FABP3	BGN	DNAJB9	FURIN	SUMF2	HSPA1A	HSPA1B
All biological processes	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Immune system process	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Circulatory system process				■										
Regulation of DNA-templated transcription	■						■						■	■
Protein folding													■	■
Lipid metabolic process						■		■						
Autophagy							■							
Inflammatory response	■		■	■		■	■				■			
Mitochondrion organization													■	■
Cytoskeleton organization							■						■	■
Chromosome segregation													■	■
Cell adhesion							■						■	■
Programmed cell death							■						■	■
mRNA metabolic process													■	■
Signaling	■					■	■			■	■		■	■
Cell differentiation			■	■		■	■	■		■			■	■
Protein catabolic process											■		■	■
Extracellular matrix process											■		■	■
Wound healing						■							■	■
Anatomic structure development						■	■		■	■	■		■	■
Cell motility	■		■				■						■	■
Protein maturation											■			
Transmembrane transport						■							■	■
Nucleobase-containing small molecule metabolic process													■	■
Membrane organization							■						■	■
Protein-containing complex assembly													■	■
Defense response to other organism	■	■	■	■		■	■						■	■
Mitotic nuclear division													■	■
Carbohydrate derivative metabolic process													■	■
Other biological process	■	■	■	■		■	■	■		■	■		■	■

**Extended Data Fig. 8 | Fig. 8. Biological processes associated with the 14 mediator proteins linking social disadvantage and hallmark-specific ARDs, as identified from the Universal Protein Resource (UniProt, a comprehensive,**

**freely accessible database of protein sequence and functional information).** UniProt descriptions of biological processes for mediator proteins of the associations between social disadvantage and hallmark-specific diseases.

## 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 for data collection.

Data analysis 'organage package' in Python (<https://github.com/hamiltonoh/organage>); RStudio (2023.03.1), including STRINGdb package, clusterProfiler package and GO-term enrichment analysis, SAS statistical software (version 9.4), and Stata (17.0). Statistical code is available in <https://github.com/rmjdmj1/SES-ageing.git> and Supplementary Information

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](#)

We used pseudonymised/anonymised individual-level data from four prospective cohort studies: the UK Biobank, the Finnish Public Sector study, the Whitehall II study, and the Atherosclerosis Risk in Communities study. The statistical code is available in <https://github.com/rmjdmj1/SES-ageing.git> and Supplementary

Information. Extensive summary data are provided in 33 Supplementary tables. Researchers registered with UK Biobank can apply for access to the individual-level data by completing an application. This must include a summary of the research plan, data-fields required, any new data or variables that will be generated, and payment to cover the incremental costs of servicing an application (<https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>). In the Finnish Public Sector study, pseudonymised individual-level questionnaire data as used in this study can be shared by request to the investigators ([jaana.e.pentti@helsinki.fi](mailto:jaana.e.pentti@helsinki.fi)). Linked health records require separate permission from the Findata, the Health and Social Data Permit Authority in Finland (<https://findata.fi/en/permits/>). Pseudonymised individual-level data from the Whitehall study are available for sharing within the scientific community. Bona fide researchers interested in accessing the pseudonymised data can apply through the Dementias Platform UK (<https://www.dementiasplatform.uk>) or the Whitehall Scientific committee (<https://www.ucl.ac.uk/epidemiology-health-care/research/epidemiology-and-publichealth/research/whitehall-ii/data-sharing>). ARIC proteomic data are available through the NHLBI Biologic Specimen and Data Repository Information Coordinating Center (<https://biolincc.nhlbi.nih.gov/studies/aric>). Additional requests for clinical or proteomic data from individual investigators may be submitted to the ARIC steering committees and will be reviewed to ensure that data can be shared without compromising patient confidentiality or breaching intellectual property restrictions. Participant-level demographic, clinical and proteomic data may be partially restricted based on previously obtained participant consent. Data-sharing restrictions may also be applied to ensure consistency with confidentiality or privacy laws and considerations (<https://sites.csc.unc.edu/aric>).

## 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

Only the term 'sex' is used in this article as we study biological feature. Sex was self-reported. We have reported data disaggregated for sex. Table 1 and Supplementary Tables S1, S5, S22, S26, and S27 provide sex-specific participant numbers and results for main analysis. No significant sex differences were observed.  
Results on social disadvantage and hallmark-specific ARDs:  
"These findings were robust, as similar patterns were observed in analyses using an alternative indicator of adult SES—low occupational position—available for FPS participants and in both men and women, although the effects were stronger in men (Supplementary Tables 4, 5)." P. 5, lines 130-132.  
Results on social disadvantage and plasma proteins:  
"Sex-stratified analyses showed a similar overall pattern of results in both men and women (Supplementary Table 22)." P. 8, lines 219-220.

### Reporting on race, ethnicity, or other socially relevant groupings

We adjust analyses for ethnicity as a potential confounder. Our topic is social disadvantage and for this reason we use education, occupational group and residential social deprivation as our main exposure variables. Results regarding social status are reported in Figures 2–4 and Supplementary Information Tables 2–33.

### Population characteristics

In addition to the above, we measured morbidity associated with cellular ageing, the main outcome of this study, plasma proteins, the main mediators in the study, and polygenic risk score for education. Mean age and proportion of men were 57 and 46% in UK Biobank, 46 and 27% in the Finnish Public Sector study, 56 and 71% in the Whitehall study, 57 and 44% in the ARIC middle-aged cohort and 76 and 57% in the ARIC old-age cohort.

### Recruitment

The UK Biobank study is a nationwide, prospective cohort study of half of a million participants aged 208 between 38 and 73 years, living in the United Kingdom. The Finnish Public Sector study and the Whitehall II study are occupational cohort studies. The prevalence and incidence of disease in occupational cohort studies are lower than in the general population, but the relationships between risk factors and disease are similar. The ARIC study is multi-site, prospective, biracial cohort study. Participants were recruited from 4 U.S. communities: Forsyth County, North Carolina, Jackson, Mississippi, Suburbs of Minneapolis, Minnesota, and Washington County, Maryland.

### Ethics oversight

The North-West Multi-Centre Research Ethics Committee, the Ethical Committee of the Helsinki and Uusimaa hospital district, the University College London Hospital Committee on the Ethics of Human Research, and institutional review boards at each ARIC participating center: University of North Carolina at Chapel Hill, Chapel Hill, NC; Wake Forest University, Winston-Salem, NC; Johns Hopkins University, Baltimore, MD; University of Minnesota, Minneapolis, MN; and University of Mississippi Medical Center, Jackson, MS.

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](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Behavioural & social sciences study design

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

### Study description

Four prospective cohort studies including quantitative data on sociodemographic characteristics, socioeconomic circumstances, plasma proteins, prevalent health conditions at follow-up and incident diseases and mortality at follow-up.

### Research sample

We used existing datasets from major well-characterised population-based and occupational cohort studies, containing detailed information on socioeconomic circumstances, health, and plasma proteins. Sample sizes varied between 6544 and 492,257

	depending on the cohort. Mean age and proportion of men were 57 and 46% in UK Biobank, 46 and 27% in the Finnish Public Sector study, 56 and 71% in the Whitehall study, 57 and 44% in the ARIC middle-aged cohort and 76 and 57% in the ARIC old-age cohort. These cohorts are not directly representative of the general adult population.
Sampling strategy	UK Biobank is based on a random sample of the general population. The Finnish Public Sector study and the Whitehall II study included an occupational population from specific workplaces. ARIC is biracial cohort with participants recruited from 4 centres. The cohort studies were selected because of their large sample size and high-quality prospective data on the exposures, biological mediators, and health outcomes of interest. No power calculations were needed given the large sample size. As indicated in our Data Availability section, these datasets are available for bona fide researchers on request.
Data collection	Data collection methods, including questionnaire surveys, clinical examinations, blood analyses, and linkage to electronic health records are described in detail in the manuscript. The authors were blinded to the data collection process.
Timing	Data collection started in 2006 (UK Biobank), 2000 (Finnish Public Sector study), 1997 (Whitehall II), and 1990 (ARIC). Continuous follow-up through linked electronic health records and repeated clinical examinations ended in 2021. There were no gaps in the follow-up.
Data exclusions	Please see Extended data figures 1 and 2. Participants were excluded from the study based on predefined criteria, including missing data on exposure or outcome, non-response, or lack of consent for health record linkage.
Non-participation	Please see Extended data figures 1 and 2. Participation in the cohort studies was voluntary. Among the invited participants (eligible population), non-participation varied between 26% in the Whitehall II study and 95% in the UK Biobank.
Randomization	In this real-life study, participants were not randomized by social disadvantage; differences in covariates were addressed through statistical adjustments. Consistent findings across diverse populations and subgroups further support the robustness of our results.

## 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 checked="" type="checkbox"/>	<input 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 checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Plants

Seed stocks	<i>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.</i>
Novel plant genotypes	<i>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.</i>
Authentication	<i>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.</i>