Research Square

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

Disease aging clocks in the plasma proteome stratify risk and track preclinical progression

Zihuai He

zihuai@stanford.edu

Stanford University https://orcid.org/0000-0002-8220-4183

Junyoung Park

Stanford University https://orcid.org/0000-0002-9098-2858

Jiawei Tan

Stanford University School of Medicine

Ru Wen

Stanford University School of Medicine

Julie Wang

Stanford University

Kejun Ying

Stanford University School of Medicine https://orcid.org/0000-0002-1791-6176

Yann Le Guen

Stanford University School of Medicine https://orcid.org/0000-0001-6649-8364

Elizabeth Mormino

Stanford University

Victor Henderson

Stanford University https://orcid.org/0000-0003-1198-9240

Tony Wyss-Coray

Stanford University School of Medicine https://orcid.org/0000-0001-5893-0831

Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10675323/v1

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

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

Disease aging clocks in the plasma proteome stratify risk and track preclinical progression

Junyoung Park[1,2], Jiawei Tan[3,4], Ru M. Wen[3], Julie Fangran Wang[5], Kejun Ying[1,6], Yann Le Guen[2], Elizabeth C. Mormino[1], Victor W. Henderson[1,7], Tony Wyss-Coray[1,6,8], Zihuai He[1,2,3*]

[1]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA 94305, USA
[2]Quantitative Sciences Unit, Division of Computational Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA
[3]Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, CA 94305, USA
[4]Department of Nephrology, Stanford School of Medicine, Stanford, CA, USA
[5]Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305, USA
[6]Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, USA
[7]Department of Epidemiology and Population Health, Stanford University School of Medicine, Stanford, CA 94305, USA
[8]The Phil and Penny Knight Initiative for Brain Resilience, Stanford University, Stanford, CA, USA

[*]Correspondence: zihuai@stanford.edu

ABSTRACT

Aging is the strongest risk factor for most chronic diseases, yet people of the same chronological age differ in how far disease processes have advanced. Existing plasma proteomic clocks estimate chronological age, or the age of a single organ, but none targets the pathway from aging to a specific disease. Here we developed disease aging clocks by deriving the plasma proteomic composite that maximizes the age-mediated component of disease risk and applied the framework to 31 age-related diseases in 44,526 UK Biobank participants. The resulting disease age gap separated 10-year incidence up to 20-fold among individuals of the same chronological age, and each additional year of gap raised the hazard of type 2 diabetes by 24%, chronic kidney disease by 21% and Alzheimer's disease (AD) by 17%. Among participants who went on to develop disease, the gap was larger closer to diagnosis in 19 of 20 diseases, indicating that it reports preclinical stage rather than fixed susceptibility. Across these 20 diseases, 56% of the gap was orthogonal to chronological and organ-specific clocks. Applied without refitting to Stanford Alzheimer's Disease Research Center participants (n=501), the AD clock was elevated in AD and Lewy body disease but not Parkinson's disease, and tracked plasma biomarkers, cognition and neuroimaging features. Disease aging clocks provide a blood-based measure of the stage of a specific aging-to-disease process years before diagnosis, with potential applications in prevention-trial enrichment and longitudinal disease monitoring.

Age is the strongest risk factor for most major chronic diseases, including heart failure, dementia, type 2 diabetes (T2D) and chronic kidney disease (CKD)[1–3]. Yet individuals of the same chronological age differ substantially in how far they have progressed toward each disease. Two people of the same age may need very different prevention or monitoring strategies. Effective prevention depends on detecting disease during the long preclinical period, when age-related molecular changes are already accumulating but clinical symptoms remain absent. Existing risk-assessment tools, including chronological age, family history and polygenic risk scores[4,5], estimate who is likely to develop disease, but none reports how far the age-related disease process has already progressed at the time of measurement. What is missing is a disease-specific, blood-based measure of preclinical disease stage.

The plasma proteome integrates molecular signals across organs and physiological systems, responds to changes in physiological state, and can be sampled repeatedly from blood[6,7]. Plasma proteomic aging clocks use these signals to estimate chronological age with high accuracy, and accelerated proteomic aging is associated with mortality, multimorbidity and a range of age-related diseases[8,9]. Organ aging signatures refined this approach by resolving biological aging anatomically, showing that different organs can age at different rates within the same individual[10–12]. They localize biological aging to specific organs but do not quantify the disease-related molecular state associated with any particular disease. Predicting chronological age is a natural starting point, but the proteins that best track chronological time are not necessarily those that best capture how aging drives a specific disease[13–17]. Disease prediction models pursue a separate goal, estimating future disease incidence from plasma proteins[18]. Because they are optimized for discrimination, they can draw on any informative signal, whether age-related disease biology, genetic or environmental influence, comorbidity or subclinical disease already underway. Existing aging clocks and disease-prediction models therefore estimate different quantities, and neither was designed to measure the age-related molecular state underlying preclinical progression toward a specific disease.

We reasoned that a clinically relevant quantity is disease age, a disease-specific plasma proteomic measure of the molecular component that links chronological aging to disease. Disease aging clocks estimate disease age from a single plasma proteomic profile and express it on the scale of chronological age. Each disease-specific model is trained to maximize the age-mediated component of disease risk captured by the plasma proteome[19]. This objective favors proteins that change with age and are also associated with the disease after accounting for age. The difference between disease age and chronological age defines the disease age gap. A positive gap indicates that the age-related molecular state for that disease is more advanced than expected for the individual's chronological age. We hypothesized that disease age would distinguish individuals of the same chronological age by their risk of a specific disease years before diagnosis.

To test this hypothesis, we developed disease aging clocks for 31 age-related diseases using plasma proteomic profiles from 44,526 UK Biobank (UKB) participants. We first evaluated whether disease age stratified future disease incidence and reflected proximity to clinical diagnosis among individuals of the same chronological age. We next assessed the transportability of the Alzheimer's disease (AD) clock by directly applying the UKB model to an independent Stanford Alzheimer's Disease Research Center (ADRC) cohort without model refitting. Finally, to evaluate cross-platform reproducibility, we independently trained and evaluated the disease aging framework in the Global Neurodegeneration Proteomics Consortium (GNPC) using SomaScan proteomic measurements. Together, these analyses evaluated disease age across prospective population, clinical, and cross-platform settings.

Disease age quantifies a disease-specific preclinical molecular state

We sought to quantify the molecular state through which chronological aging contributes to individual diseases rather than biological aging itself. For each disease, we identified the plasma-protein composite that maximizes the indirect effect of chronological age on disease risk using MaxIE (maximum indirect effect)[19], a closed-form optimization framework (Fig. 1a; Methods). The resulting linear protein composite defines a disease aging clock, and its prediction, expressed in chronological-age units, defines disease age.

We applied this framework to 2,916 plasma proteins measured in 44,526 UKB participants across 31 age-related diseases (Fig. 1b). We first tested whether each disease exhibited evidence of an age-mediated proteomic pathway. Twenty diseases showed nominally significant proteome-mediated aging signals and were carried forward for disease age model construction. Disease-specific protein weights were estimated using a geographically defined training set (n = 21,504) and evaluated in an independent held-out UKB test set (n = 23,022). We then evaluated disease age across complementary settings. Within the UKB, we tested prospective disease stratification and whether larger baseline disease age gaps were associated with shorter time to clinical diagnosis. For AD, we directly applied the UKB model to an independent Stanford ADRC cohort without model refitting and further independently trained and evaluated the disease aging framework in the GNPC SomaScan cohort to assess reproducibility across proteomic platforms.

The plasma proteome shows evidence of aging-to-disease mediation across 20 diseases

Constructing a disease aging clock requires that the plasma proteome carries an aging-to-disease signal. We first applied the cosine test (Methods) to determine whether proteomic changes associated with chronological age were aligned with those associated with each disease. Fifteen of 31 diseases reached Bonferroni significance ($p < 0.0016$) and 20 reached nominal significance ($p < 0.05$; Fig. 2a; Supplementary Table 1). The strongest evidence was seen for cardiovascular disease (hypertensive disease, $p = 5.1 \times 10^{-13}$; heart failure, $p = 8.0 \times 10^{-13}$; cerebrovascular disease, $p = 1.2 \times 10^{-11}$; ischemic heart disease, $p = 4.1 \times 10^{-10}$), CKD ($p = 5.1 \times 10^{-10}$) and neurodegenerative conditions (all-cause dementia, $p = 8.9 \times 10^{-10}$; AD, $p = 9.3 \times 10^{-9}$). Eight further diseases passed Bonferroni, spanning musculoskeletal (osteoarthritis, $p = 1.5 \times 10^{-8}$; osteoporosis, $p = 6.4 \times 10^{-6}$), respiratory (chronic obstructive pulmonary disease (COPD)/emphysema, $p = 4.4 \times 10^{-8}$), cardiac (atrial fibrillation, $p = 6.1 \times 10^{-7}$), neurological (vascular dementia, $p = 5.2 \times 10^{-5}$), metabolic (type 2 diabetes, $p = 3.0 \times 10^{-4}$), and cancer (lung, $p = 3.3 \times 10^{-5}$; esophageal, $p = 4.1 \times 10^{-4}$) categories. By contrast, six cancers showed no detectable mediation: breast, brain, ovarian, colorectal, prostate, and pancreatic cancer. For breast (n = 805), prostate (n = 739), and colorectal cancer (n = 479), case counts exceeded most Bonferroni-pass diseases, so the null result is unlikely to reflect low power. The cosine test is therefore positive only when aging affects risk through the proteome. Age-associated mechanisms in cancer such as somatic mutation accumulation, hormonal changes, and microenvironment remodeling may not produce a coordinated proteomic signature in this middle-aged cohort (mean 56.8 years).

For each of the 20 nominally significant diseases, we estimated disease-specific protein weights in the training set, calibrated the resulting disease age measure to year-equivalent disease age units and evaluated the frozen models in the independent test set. Calibration to chronological age units was accurate. For example, the T2D model showed a correlation of 0.89 with chronological age and a slope of 0.98 in the test set (Fig. 2b). Accuracy was consistent across all 20 clocks (test-set

r 0.89–0.90, slope 0.97–0.98; Supplementary Table 2). Sex-specific calibration centered the disease age gap at zero in both women and men, with a standard deviation of 4.1 years for T2D (Fig. 2c). Across all 20 conditions, prevalent cases had higher disease age gaps than same age controls, with median differences ranging from 1.2 to 4.6 years (median, 1.9 years; Fig. 2d). These cross-sectional differences provided initial evidence that disease age captures disease-relevant molecular variation before prospective evaluation.

Disease age captures disease-relevant aging biology beyond organ aging signatures

A key question was whether disease age simply recapitulates information already contained in existing organ aging signatures or instead captures additional disease-relevant aging biology. We compared disease aging clocks with the plasma proteomic organ aging signatures developed by Oh et al.[10], the closest existing framework because they resolve biological aging at the level of individual organs and predict organ-specific disease outcomes. As these signatures have already been shown to provide greater disease relevance than conventional organism-wide aging models, we focused the comparison on organ aging signatures.

We first quantified how much disease age variation was shared with existing proteomic aging signatures. In the held-out test set, we decomposed each clock's variance into that explained by chronological age, additional variation shared with 20 organ aging clocks reported by Oh et al., and the remaining component not accounted for by these signatures. Chronological age explained, on average, 79.7% of the variation in disease age across the 20 diseases. The remaining 20.3% represents age-independent variation among individuals of the same chronological age (the disease age gap). Of this gap, all existing clocks collectively explained on average 8.8 percentage points (=43.6% of the gap), while an additional 11.5 percentage points (=56.4%) remained orthogonal to every existing clock (Extended Data Fig. 1). This disease-relevant aging signal is not captured by current proteomic aging clocks.

We next compared disease age with organ aging signatures using an age-mediated disease association metric that quantifies how strongly each signature captures the component of disease associated with aging. Disease age showed stronger age-mediated disease associations than the best-matched organ aging signature in 17 of the 20 diseases (Fig. 3a), with the largest differences observed for T2D, CKD and COPD. Disease age also remained consistently calibrated to chronological age across diseases (r=0.89-0.90), whereas organ aging signatures showed substantially greater variability (r=0.34–0.72; Fig. 3b). These differences translated into improved disease discrimination. Disease age achieved higher classification performance than chronological age alone in all 20 diseases (Fig. 3c) and exceeded the best-performing organ aging signature in 16 of 20 diseases (Fig. 3d). The four exceptions comprised frontotemporal dementia (FTD), vascular dementia, amyotrophic lateral sclerosis (ALS) and esophageal cancer. FTD and ALS are known to involve protein-homeostasis pathways that are only partly mediated by aging[20]. Accordingly, organ aging signatures may capture organ-specific biology beyond the age-mediated component isolated by disease age. The same comparison for the 11 diseases that did not reach nominal significance in the cosine test is in Supplementary Fig. 1.

Disease age stratifies current and future disease risk

Disease age gap was strongly associated with prevalent disease across the 20 nominally significant diseases in cross-sectional analysis (all Bonferroni-significant after adjustment for chronological age; Fig. 4a). Among cardiometabolic diseases, the odds ratio per year of gap was OR = 1.32 (95% confidence interval, 1.30 – 1.33) for T2D, 1.24 (1.22 – 1.25) for CKD; among

musculoskeletal conditions, 1.23 (1.21–1.25) for rheumatoid arthritis; among hematological conditions, 1.25 (1.21–1.29) for leukemia; and among neurodegenerative diseases, 1.19 (1.16–1.22) for AD. Prospective Cox proportional hazards analysis excluding prevalent cases (median follow-up 10.4 years; maximum 17.6) yielded consistent associations, with no evidence of survival bias (Fig. 4b). Each year of disease age gap increased the hazard ratio for T2D by 24% (HR = 1.24, 1.23–1.26), COPD by 23% (1.23, 1.21–1.24), CKD by 21% (1.21, 1.20–1.22), AD by 17% (1.17, 1.15–1.20), rheumatoid arthritis by 17% (1.17, 1.14–1.19), and heart failure by 15% (1.15, 1.14 – 1.17). Associations were robust across disease categories, from cardiovascular and metabolic to neurodegenerative and musculoskeletal conditions. Results for the other 11 diseases are in Supplementary Fig. 2.

Among individuals of the same chronological age, disease age gap stratified future disease incidence by up to 20-fold over 10 years (Fig. 4c). Participants were classified into the lowest decile, middle 80%, and highest decile of the disease age gap. Because disease age gap is orthogonal to chronological age by construction, these groups were matched for chronological age and differed primarily in their disease-related molecular state. Ten-year cumulative incidence differed markedly between the highest- and lowest-gap groups, reaching 20-fold for AD (3.8% versus 0.2%), 16-fold for T2D (29.2% versus 1.8%), 10-fold for CKD (15.7% versus 1.5%), 7-fold for heart failure (12.6% versus 1.9%), and 6-fold for rheumatoid arthritis (5.0% versus 0.9%) (all 20 diseases are shown in Supplementary Fig. 3). Although these groups differed by only approximately eight disease age years, this modest difference identified individuals with strikingly different future disease incidence. In contrast, full disease age incorporates chronological age to estimate absolute disease risk. Stratification by full disease age produced even greater separation of cumulative incidence, with more than 25-fold differences for AD, heart failure and CKD over 15 years (Fig. 4d; all 20 diseases are shown in Supplementary Fig. 4). Thus, disease age gap provides age-independent risk stratification, whereas full disease age translates this information into absolute disease risk.

The disease age gap tracks preclinical disease progression

Disease age reflected not only future disease risk but also proximity to clinical diagnosis. A conventional risk score can identify individuals who will eventually develop disease but remain unchanged throughout the preclinical period. By contrast, if disease age captures disease stage, individuals closer to clinical onset should already exhibit larger disease age gaps at baseline. To distinguish disease stage from disease susceptibility, we restricted the analysis to incident cases only and tested whether baseline disease age gap was associated with time to first diagnosis (Fig. 4e; Extended Data Fig. 2). Because every participant in this analysis eventually developed the disease, the analysis evaluated differences in disease stage rather than differences in lifetime susceptibility.

Among all 20 nominally significant diseases (all with at least 50 incident cases), baseline disease age gap systematically increased as clinical diagnosis approached. Linear regression showed a negative association between disease age gap and time to diagnosis in 19 of 20 diseases, with 17 reaching nominal significance ($P < 0.05$) and 15 remaining significant after Bonferroni correction ($P < 0.0025$). The three exceptions (esophageal cancer, FTD and lung cancer) did not reach significance, consistent with their relatively small numbers of incident cases. The strongest association was observed for CKD, where participants diagnosed within three years had a mean disease age gap of +11.0 years, compared with +3.8 years among those diagnosed 7–10 years later. Similar gradients were observed for T2D (+6.2 versus +3.1 years) and rheumatoid arthritis (+5.7

versus +2.0 years), and the same pattern extended across cardiovascular, metabolic, neurodegenerative and musculoskeletal diseases. This inverse relationship between gap size and time to onset is inconsistent with a static susceptibility score, which would predict who develops disease but not modulate with proximity to onset. It is consistent with the gap reflecting how far a subclinical disease process has advanced at the time of measurement, a biological stage rather than a fixed trait.

Disease aging clock weights encode shared and disease-specific biological structure

Each disease aging clock was defined by a vector of 2,916 protein weights, representing the direction and relative magnitude of each protein's contribution to the disease age estimate. Correlations among these weight vectors grouped the clocks by clinical category, even though each was trained independently (Fig. 5a). Neurological diseases (AD, all-cause dementia, vascular dementia, FTD, ALS) formed one cluster ($r > 0.6$ within group). Cardiovascular diseases (hypertensive disease, ischemic heart disease, heart failure, atrial fibrillation) formed a second, and musculoskeletal diseases a third. Uniform manifold approximation and projection embedding of the weight vectors confirms this structure (Fig. 5b). Two cancers, lung and esophageal, did not group with each other or with any other category.

Pathway enrichment on the weight vectors separated a shared component from disease-specific ones (fgsea; Fig. 5c). The first principal component (PC1) of the weight matrix accounted for 53% of cross-disease weight variance and was enriched for regulation of neurogenesis (false discovery rate (FDR) < 0.005), elastic fibre formation (FDR < 0.01) and epithelial–mesenchymal transition (FDR < 0.05). Beyond this shared component, the AD clock was enriched for CSF1/M-CSF signaling in myeloid cells and CD28 co-stimulation, together with positive regulation of neurogenesis, nervous system development and small-GTPase-mediated signal transduction (all FDR < 0.05), consistent with reported immune contributions to AD[21,22]. The T2D clock was enriched for epithelial–mesenchymal transition and extracellular matrix organization (FDR < 0.01) and for enzyme-linked receptor protein signaling (FDR < 0.05). The heart failure clock was enriched for cytosolic DNA sensing, enzyme-linked receptor protein signaling and endocytosis (FDR < 0.05). The CKD clock showed the most enrichments, led by epithelial–mesenchymal transition (FDR < 0.001) and ERBB signaling (FDR < 0.005), with eight further pathways at FDR < 0.05 spanning growth-factor signaling, matrix organization and platelet activation. No pathway survived FDR correction for the rheumatoid arthritis clock. Enrichment results for the shared component and all 20 disease aging clocks are provided in Supplementary Table 3.

Because the biological analyses above rely on the interpretability of the linear weights, we also tested whether more flexible nonlinear representations improved performance. Across the two diseases examined, AD and T2D, autoencoders, tree ensembles and kernel methods did not outperform the linear model (Supplementary Note; Supplementary Table 4). The linear weights therefore provide an additive and directly quantifiable protein-level decomposition of disease-centered aging biology.

The AD disease age generalizes to an independent clinical cohort and remains disease-specific: the Stanford ADRC

All preceding analyses were performed in the UKB, a population-based cohort of middle-aged adults. To evaluate whether the disease age generalizes to an independent clinical cohort, we applied the UKB-trained AD model, without refitting, to Stanford ADRC participants (n=501; mean age 70.7 $\pm$8.2 years; plasma Olink Explore 3k; see Methods for diagnostic composition)[12].

Because follow-up in the ADRC was short (median=1 year), this analysis assessed biological validity rather than prospective prediction. After correcting for cohort-level platform differences using intercept calibration, AD disease age remained well calibrated in healthy controls (HC) despite an approximately 14-year difference in mean chronological age between cohorts (mean disease age gap: -1.28 years in ADRC versus −0.05 years in UKB; Fig. 6a).

Disease age remained highly disease specific in the clinical cohort. Compared with HCs, the AD disease age gap was selectively elevated in AD (mean +6.3 years above HC; two-sided Wilcoxon $p = 5.9 \times 10^{-8}$; Fig. 6a), whereas Parkinson's disease (PD; +0.4 years), PD with mild cognitive impairment (PD-MCI; +0.9 years) and mild cognitive impairment (MCI) alone (+0.7 years) showed no significant elevation. This pattern supports that the AD disease age is not a generic marker of neurodegeneration or cognitive decline but is more closely aligned with the amyloid/tau pathological axis that underlies AD. Lewy body dementia exhibited a disease age gap comparable to AD (+6.5 years; $p = 1.7 \times 10^{-3}$), consistent with the frequent coexistence of AD-related amyloid and tau pathology in this disorder. Applied to the same participants, the UKB-derived PD clock separated PD (+3.7 years, $p = 1.2 \times 10^{-3}$) and PD-MCI (+3.7 years, $p = 1.6 \times 10^{-3}$) from HC (Fig. 6b). Its elevation in AD was smaller and weaker (+2.7 years, $p = 0.023$), indicating that disease age reflects disease-specific molecular biology rather than neurodegeneration or cognitive impairment in general.

The AD age gap also stratified diagnostic probability across cohorts. In both the Stanford ADRC and an age-matched subset of the UKB test set, the prevalence of AD increased monotonically with increasing disease age gap thresholds (Fig. 6c). In the ADRC, AD prevalence increased from approximately 20% to more than 60% across the evaluated threshold range, with a similar enrichment pattern in the UKB despite an approximately fourfold lower overall disease prevalence.

The AD age gap aligns with plasma biomarkers, cognitive measures and neuroimaging features of AD

To characterize the clinical and biological correlates of the AD age gap, we examined associations with plasma biomarkers, cognitive measures and structural magnetic resonance imaging (MRI) phenotypes using linear mixed models adjusted for age and sex. Analyses were performed separately in the HC only sample (n=183–226), the combined HC and MCI sample (n=245–313), and the full sample comprising HC, MCI and AD (n=300–380). Benjamini–Hochberg FDR correction was applied jointly across all 20 outcomes (5 plasma markers, 9 cognitive scores, 6 structural MRI measures) within each sample group. Figure 6d presents 12 clinically representative outcomes in the full sample, spanning AD plasma biomarkers, global cognition and clinical severity, episodic memory and characteristic structural neurodegeneration. Results for all 20 outcomes across all three subsamples are provided in Supplementary Table 5.

Across the full sample, the AD age gap was positively associated with four of five plasma AD markers: phospho-tau (pTau) 217, neurofilament light chain (NfL), pTau 181, and glial fibrillary acidic protein (GFAP) (all FDR $q < 10^{-4}$; Fig. 6d). The A$\beta_{42/40}$ ratio showed no significant association ($p = 0.29$), consistent with amyloid accumulation plateauing early in the disease cascade while the age gap continues to reflect downstream neurodegeneration. pTau 217 and NfL were already nominally associated in the HC only sample ($p < 0.01$), indicating that the age gap captures subclinical neurodegeneration before cognitive symptoms emerge. Cognitive associations were broad, spanning global cognition (Montreal Cognitive Assessment), episodic memory, and clinical severity (Clinical Dementia Rating Sum of Boxes; all FDR $q < 0.01$). Structural MRI associations included smaller hippocampal and amygdala volumes and lower mean cortical

thickness (all FDR $q < 0.05$). Associations generally strengthened as participants with progressively greater clinical impairment were included, consistent with the age gap tracking cumulative disease burden rather than a static pathological marker.

We next examined the regional neuroimaging pattern associated with the AD age gap in both the combined HC and MCI sample and the full sample, with the full-sample results presented in Fig. 6e and the corresponding HC and MCI results provided in Supplementary Fig. 5 (full results are in Supplementary Table 6). In the full imaging sample (n = 100), a higher AD age gap was associated with cortical thinning in AD-vulnerable regions. The association in the fusiform gyrus remained significant after FDR correction ($\beta = -0.33$, $q = 0.02$), while additional temporoparietal and occipitotemporal regions showed nominal associations. Regional tau positron emission tomography (PET) analyses showed nominal associations between a higher gap and greater tau retention in the posterior cingulate, precuneus and isthmus cingulate, although none remained significant after FDR correction. No significant regional associations were observed with amyloid PET.

An independent SomaScan clock reproduces the AD age signal in GNPC

To test the approach on a second proteomic platform, we independently retrained the AD disease age model in the GNPC[23], a multi-site SomaScan cohort comprising 13,388 participants from 17 contributing centers. We independently trained a new AD disease age model using 2,684 SomaScan aptamers targeting proteins shared with the Olink panel and evaluated performance using site-grouped five-fold cross-fitting, ensuring that each participant was evaluated using weights estimated without data from their own site.

The plasma proteome again supported an age-mediated pathway linking chronological age to AD (cosine test, $P=6.6 \times 10^{-6}$). In out-of-fold predictions, the calibrated AD disease age gap was elevated in AD relative to HC by a median of 3.3 years, with MCI showing an intermediate gap of 1.0 years (Extended Data Fig. 3a). In ten of 17 contributing sites with at least 10 AD cases and 10 controls, leave-one-site-out analysis gave AD-minus-control mean differences of 3.21 to 4.17 years (Extended Data Fig. 3b). Consistent with the UKB and Stanford ADRC analyses, the prevalence of AD increased monotonically with increasing disease age gap thresholds (Extended Data Fig. 3c). Independent recovery of the AD disease-age signal in GNPC using SomaScan shows that the framework is not specific to a single cohort or proteomic platform.

DISCUSSION

A single plasma proteomic measurement, obtained before clinical disease onset, can provide information not only about future disease risk, but also about how far the age-related disease process had already advanced. Many established risk-stratification approaches are designed mainly to identify who will develop disease and are less informative about the state of the disease process at the time of assessment. Disease age was designed to quantify this disease-specific molecular state on an age-equivalent scale. By restricting the analysis to participants who subsequently developed the disease, we asked whether baseline disease age varied according to how soon diagnosis occurred, rather than whether disease occurred at all. Across 19 of 20 conditions, individuals closer to diagnosis had larger baseline disease age gaps. This pattern is consistent with disease age providing a molecular measure of the preclinical disease process beyond susceptibility alone.

In the UKB, disease age prospectively stratified incident disease across 20 diseases, separated 10-year cumulative incidence by up to 20-fold among individuals of the same chronological age and

was higher among future cases closer to diagnosis. The AD age was then evaluated in two independent settings. When the UKB–derived AD model was applied without refitting to the Stanford ADRC cohort, the AD age gap was selectively elevated in AD and associated with plasma biomarkers, cognitive function and neuroimaging features of AD pathology, supporting its transportability to an independent clinical cohort. The AD disease age signal was also reproduced in GNPC using a model trained and evaluated independently on the SomaScan platform, supporting reproducibility of the disease age framework on a different proteomic platform.

Proteomic aging clocks estimate general biological aging, and organ aging resolves this variation across individual organs, but neither is defined in relation to a specific disease. Disease-prediction models pursue a different objective, selecting whatever proteins best discriminate future cases, regardless of whether those proteins reflect aging. Disease age isolates the age-related component of the proteome specific to a given disease, quantifying disease-specific aging rather than general aging or overall risk. Existing proteomic aging signatures, including organ-specific clocks, accounted for 43.6% of the age-independent variance in disease age. The remaining 56.4% was information they did not capture. So, disease age acne be used alongside existing clocks. Its weights are selected for the age-related proteomic component associated with a specific disease.

Each clock was trained independently, but the weight vectors clustered by clinical category, with the neurological clocks correlating at $r > 0.6$. A shared tissue-remodeling component explained approximately half of the variation across clocks, whereas disease-specific weights recovered established biology for AD, T2D and CKD. Not every disease yielded a disease age signal. Six cancers showed no detectable aging-to-disease alignment despite case numbers exceeding those of most diseases that passed Bonferroni correction. Although an existing proteomic age-prediction clock has been associated with cancer incidence[8], association with incidence does not imply an age-mediated proteomic pathway. The age-associated mechanisms that dominate cancer risk in midlife, including somatic mutation accumulation and tumor microenvironment remodeling[24], need not produce a coordinated plasma signal in this middle-aged cohort (mean age: 56.8 years).

This study has several limitations. First, the gradient between disease age gap and time to diagnosis was observed across individuals rather than within them, and longitudinal proteomics will be needed to test whether a person's own gap increases as onset approaches. Second, the observational design cannot establish causation, and the weights are optimized for mediation rather than estimating the causal effect of individual proteins. Third, external biological validation with deep biomarkers and imaging is available only for the AD age, and comparable validation is needed for the others.

Disease age is defined by optimizing the age-mediated component of a specific disease and is not trained to predict chronological age alone. The resulting measure distinguishes individuals of the same chronological age by their disease-related molecular state. The mediation criterion is not specific to plasma proteomics and could be extended to other molecular modalities and modeling frameworks. Measures constructed in this way may help identify individuals closer to clinical onset and enrich prevention trials. More broadly, they provide a framework for defining disease age according to the molecular pathways that link aging to disease.

References

1. Partridge, L., Deelen, J. & Slagboom, P. E. Facing up to the global challenges of ageing. *Nature* **561**, 45–56 (2018).
2. Kennedy, B. K. et al. Geroscience: linking aging to chronic disease. *Cell* **159**, 709–713 (2014).
3. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. Hallmarks of aging: an expanding universe. *Cell* **186**, 243–278 (2023).
4. Khera, A. V. et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
5. Inouye, M. et al. Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *J. Am. Coll. Cardiol.* **72**, 1883–1893 (2018).
6. Lehallier, B. et al. Undulating changes in human plasma proteome profiles across the lifespan. *Nat. Med.* **25**, 1843–1850 (2019).
7. Tanaka, T. et al. Plasma proteomic signature of age in healthy humans. *Aging Cell* **17**, e12799 (2018).
8. Argentieri, M. A. et al. Proteomic aging clock predicts mortality and risk of common age-related diseases in diverse populations. *Nat. Med.* **30**, 2450–2460 (2024).
9. Rutledge, J., Oh, H. & Wyss-Coray, T. Measuring biological age using omics data. *Nat. Rev. Genet.* **23**, 715–727 (2022).
10. Oh, H. S.-H. et al. Organ aging signatures in the plasma proteome track health and disease. *Nature* **624**, 164–172 (2023).
11. Wang, Y. et al. Organ-specific proteomic aging clocks predict disease and longevity across diverse populations. *Nat. Aging* **6**, 162–180 (2026).
12. Oh, H. S.-H. et al. Plasma proteomics links brain and immune system aging with healthspan and longevity. *Nat. Med.* **31**, 2703–2711 (2025).
13. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
14. Hannum, G. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).
15. Levine, M. E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
16. Lu, A. T. et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303–327 (2019).
17. Ying, K. et al. Causality-enriched epigenetic age uncouples damage and adaptation. *Nat. Aging* **4**, 231–246 (2024).
18. Carrasco-Zanini, J. et al. Proteomic signatures improve risk prediction for common and rare diseases. *Nat. Med.* **30**, 2489–2498 (2024).
19. He, Z. Learning the optimal composite mediator: closed-form solution and inference. Preprint at https://arxiv.org/abs/2603.13762 (2026).
20. Ling, S.-C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416–438 (2013).
21. Heneka, M. T. et al. Neuroinflammation in Alzheimer disease. *Nat. Rev. Immunol.* **25**, 321–352 (2025).
22. Bellenguez, C. et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat. Genet.* **54**, 412–436 (2022).

23. Imam, F. et al. The Global Neurodegeneration Proteomics Consortium: biomarker and drug target discovery for common neurodegenerative diseases and aging. *Nat. Med.* **31**, 2556–2566 (2025).
24. Hanahan, D. Hallmarks of cancer: new dimensions. *Cancer Discov.* **12**, 31–46 (2022).
25. 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).
26. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
27. Assarsson, E. et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One* **9**, e95192 (2014).
28. Wik, L. et al. Proximity extension assay in combination with next-generation sequencing for high-throughput proteome-wide analysis. *Mol. Cell Proteomics* **20**, 100168 (2021).
29. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
30. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
31. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at https://doi.org/10.1101/060012 (2021).
32. Liberzon, A. et al. The Molecular Signatures Database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
33. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
34. Geurts, P., Ernst, D. & Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **63**, 3–42 (2006).
35. Schölkopf, B. & Smola, A. J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond* (MIT Press, 2002).
36. Hinton, G. E. & Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science* **313**, 504–507 (2006).
37. Kingma, D. P. & Welling, M. Auto-encoding variational Bayes. Preprint at https://arxiv.org/abs/1312.6114 (2013).
38. Ke, G. et al. LightGBM: a highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems 30*, 3146–3154 (2017).
39. Nasreddine, Z. S. et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53**, 695–699 (2005).
40. Morris, J. C. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* **43**, 2412–2414 (1993).
41. Desikan, R. S. et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).
42. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–781 (2012).
43. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
44. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).
45. Mowinckel, A. M. & Vidal-Piñeiro, D. Visualization of brain statistics with R-packages ggseg and ggseg3d. *Adv. Methods Pract. Psychol. Sci.* **3**, 466–483 (2020).

46. Gold, L. et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* **5**, e15004 (2010).
47. UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res. 53, D609–D617 (2025).
48. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
49. Nygaard, V., Rødland, E. A. & Hovig, E. Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. *Biostatistics* **17**, 29–39 (2016).
50. Rahimi, A. & Recht, B. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems 20*, 1177–1184 (2007).
51. Williams, C. K. I. & Seeger, M. Using the Nyström method to speed up kernel machines. In *Advances in Neural Information Processing Systems 13*, 682–688 (2000).
52. 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).
53. Folstein, M. F., Folstein, S. E. & McHugh, P. R. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198 (1975).

METHODS

Study population and proteomic measurements.

We used data from the UK Biobank Pharma Proteomics Project (UKB-PPP), a substudy of the UK Biobank (UKB) prospective cohort [25,26]. The analytic sample comprised 44,526 participants with complete proteomic and phenotypic data (mean age 56.8 years, s.d. 8.2; 54% female; age range 39–70 years at recruitment). Plasma proteins were measured using the Olink Explore 3072 proximity extension assay, which quantifies 2,923 proteins on 15 panels using paired antibody proximity ligation followed by next-generation sequencing [27,28]. After standard Olink quality control (removing proteins with >10% below the limit of detection or with intra-assay coefficient of variation >20%), 2,916 proteins were retained. Protein levels were standardized to mean zero and unit variance using the training-set mean and standard deviation for each protein, following the approach of Oh et al. [10]. All normalization parameters were estimated exclusively on the training set and applied without modification to the held-out test set, preventing any data leakage from test-set observations into the normalization step. All participants provided written informed consent; the UK Biobank has ethics approval from the North West Multi-center Research Ethics Committee (reference 11/NW/0382).

Disease phenotype definitions.

Thirty-one age-related disease phenotypes were defined from International Classification of Diseases (ICD)-10 codes recorded in linked UKB hospital episode statistics (HES), death registers, and primary care data, using the earliest recorded occurrence across all sources; full ICD-10 code mappings are provided in Supplementary Table 7. Diseases spanned five categories: cardiometabolic and respiratory disorders (type 2 diabetes, ischemic heart disease, hypertensive disease, heart failure, atrial fibrillation, chronic kidney disease, chronic liver disease, cerebrovascular disease, chronic obstructive pulmonary disease (COPD)/emphysema), neurodegenerative and neurological diseases (Alzheimer's disease, all-cause dementia, vascular dementia, frontotemporal dementia, Parkinson's disease, amyotrophic lateral sclerosis), psychiatric conditions (depression), cancers (colorectal, lung, breast, prostate, leukemia, non-Hodgkin lymphoma, esophageal, liver, pancreatic, brain, and ovarian), and musculoskeletal and ophthalmic conditions (rheumatoid arthritis, osteoporosis, osteoarthritis, macular degeneration). For cross-sectional analyses, cases were anyone with the condition ever recorded in UKB health records. For prospective analyses, only incident cases (first recorded event strictly after the proteomics assessment date) were included. Participants with a prior diagnosis before proteome assessment were excluded.

The cosine test for composite mediation.

The cosine test evaluates the null hypothesis H_0 that no linear combination of proteins mediates the effect of chronological age A on disease outcome Y [19]. Let $\boldsymbol{X} \in \mathbb{R}^{n\times p}$ denote the matrix of standardized protein levels, $\boldsymbol{a} = \boldsymbol{X}^\top \boldsymbol{A}$ the protein–age association vector, and $\boldsymbol{z} = \boldsymbol{X}^\top \boldsymbol{Y} - (\boldsymbol{A}^\top \boldsymbol{Y}/\| \boldsymbol{A} \|^2)\, \boldsymbol{a}$ the age-residualized protein–disease association vector. The age-residualized Gram matrix is $\boldsymbol{V} = \boldsymbol{X}^\top \boldsymbol{X} - \boldsymbol{a}\boldsymbol{a}^\top/\| \boldsymbol{A} \|^2$. Defining whitened path vectors $\boldsymbol{p} = \boldsymbol{V}^{-1}\boldsymbol{a}$ and $\boldsymbol{q} = \boldsymbol{V}^{-1}\boldsymbol{z}$, the test statistic is

$$T \;=\; \cos\varphi \cdot \sqrt{\frac{p-1}{1-\cos^2\varphi}}, \qquad \cos\varphi \;=\; \frac{\boldsymbol{a}^\top \boldsymbol{q}}{\sqrt{\boldsymbol{a}^\top \boldsymbol{p}}\,\sqrt{\boldsymbol{z}^\top \boldsymbol{q}}},$$

which follows $t(p-1) = t(2{,}915)$ under H_0 [19]. p-values are two-sided, consistent with the original derivation of the test [19], which does not assume a directional alternative. The cosine is a by-product of the solver described below and requires no additional computation. Bonferroni correction for 31 simultaneous tests was applied throughout (threshold $p < 0.05/31 = 0.0016$).

MaxIE: constructing disease aging clocks.

The general problem of disease aging clock construction is to find a mediator $M(\boldsymbol{X}^*)$ that maximally transmits the effect of chronological age on disease, where $\boldsymbol{X}^*$ is any feature representation of the molecular measurements $\boldsymbol{X}$: $\boldsymbol{X}^*$ may be the raw measurements themselves, a nonlinear transformation, a set of biology-derived pathway scores, or any other encoding. MaxIE solves this problem for the class of linear mediators $M(\boldsymbol{X}^*) = \boldsymbol{X}^*\boldsymbol{w}$, finding the weight vector that maximizes the natural indirect effect, the product of the ordinary least squares (OLS) path coefficients:

$$\boldsymbol{w}^* \;=\; \arg\max_{\boldsymbol{w}\in\mathbb{R}^p} \; \hat{\alpha}(\boldsymbol{w})\,\hat{\beta}(\boldsymbol{w}),$$

where $\hat{\alpha}(\boldsymbol{w}) = \boldsymbol{A}^\top(\boldsymbol{X}^*\boldsymbol{w})/\| \boldsymbol{A} \|^2$ is the OLS coefficient of the mediator on age, and $\hat{\beta}(\boldsymbol{w}) = (\boldsymbol{Z}\boldsymbol{w})^\top\boldsymbol{Y}/\| \boldsymbol{Z}\boldsymbol{w} \|^2$ is the OLS coefficient of disease on the age-residualized mediator, where $\boldsymbol{Z} = Q_A\boldsymbol{X}^*$, $Q_A = I - \boldsymbol{A}\boldsymbol{A}^\top/\| \boldsymbol{A} \|^2$. The closed-form solution is the bisector of the whitened path vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ in the $\boldsymbol{V}$-inner product, with $\boldsymbol{w}^* \propto \boldsymbol{p}/\| \boldsymbol{p} \|_{\boldsymbol{V}} + \boldsymbol{q}/\| \boldsymbol{q} \|_{\boldsymbol{V}}$, computable at cost $\mathcal{O}(np + p^3/3)$ via a single Cholesky factorisation of $\boldsymbol{V}$. In the default setting $\boldsymbol{X}^* = \boldsymbol{X}$ (raw standardized protein levels); MaxIE weights were estimated on the training set for all 31 diseases independently and applied to test-set proteins to generate out-of-sample composite scores. The scale-free mediation index used for clock comparison is $f^* = \mathrm{cor}(\boldsymbol{X}^*\boldsymbol{w}^*, \boldsymbol{A}) \cdot \mathrm{cor}(\boldsymbol{Z}\boldsymbol{w}^*, \boldsymbol{Y})$, which is bounded in $[-1,1]$ and comparable across clocks of different construction. A full derivation of the closed-form solution, the asymptotic distribution of the cosine statistic, and simulation studies of its size and power are given in [19].

Train–test split.

To prevent overfitting and ensure that all reported performance metrics are out-of-sample, participants were partitioned by UKB assessment center into a training set ($n = 21{,}504$, 10 centers) and a held-out test set ($n = 23{,}022$, 12 centers), ensuring geographic independence between parameter estimation and evaluation. All disease clock weights, calibration coefficients, regression coefficients, and clock-comparison metrics were estimated exclusively on the training set. Every metric reported in Results was computed on the held-out test set. The sole exception is the cosine test, which evaluates a global geometric property of the protein space and uses all 44,526 samples to maximize power. Sex is handled during clock construction: the sex-personalized calibration centers the disease age gap at zero for both males and females, so the gap Δ is sex-free by construction. Downstream regression models (logistic, Cox) therefore require no sex covariate; the only adjustment is for chronological age A, which is equivalent to modeling Δ directly.

Calibration to year-equivalent disease age.

Raw MaxIE composite scores $M = \boldsymbol{X}\boldsymbol{w}^*$ are in arbitrary units. To express disease age in years and make the clock interpretable as a biological analogue of chronological age, we calibrated each clock by linear regression on the training set: $M = \beta_0 + \beta_1 A + \beta_2 S + \varepsilon$, where A is chronological age and S is sex $(0/1)$. The disease age gap is $\Delta = (M - \widehat{M})/\beta_1$, in years, and disease age is bio age $= A + \Delta$. By construction, Δ has mean zero and is uncorrelated with chronological age within the training set. Sex-personalized calibration used a single regression including sex as a

covariate ($\beta_2 S$) to eliminate systematic sex bias in the gap; both male and female gap distributions were centered at zero after correction.

Prospective survival and cross-sectional analyses.

For each disease with time-to-event data, a Cox proportional hazards model was fitted on the test set after excluding prevalent cases: $h(t) = h_0(t)\exp(\gamma_1 \text{ bio age} + \gamma_2 A)$. Adjusting for chronological age A alongside disease age is equivalent to evaluating Δ, since bio age $= A + \Delta$; the coefficient γ_1 thus captures the effect of the age gap per year, independent of chronological age. For cross-sectional analysis, logistic regression of the same form was fitted on prevalent cases and controls in the test set. Kaplan–Meier cumulative incidence curves with 95% confidence intervals (exponential Greenwood formula) were computed for two stratifications of test-set participants. The first used the sex-calibrated disease age gap Δ (Young: bottom 10%; Normal: middle 80%; Old: top 10%), which compares biological aging rate among same-age peers. The second used full disease age (raw MaxIE score), which provides absolute rather than relative risk. Follow-up was measured from the proteomics assessment date to the first recorded disease event, death, or administrative censoring, whichever occurred first; median follow-up was 10.4 years. Cox models and Kaplan–Meier estimates were computed with lifelines 0.29, the logistic models in SciPy 1.12, and areas under the curve (AUC) with scikit-learn 1.4.

Age gap and time to disease onset.

This analysis was restricted to participants who developed incident disease during follow-up excluding disease-free controls. For each such participant, the time to onset was defined as the interval from the proteomics assessment to the first recorded disease event. For each of the 20 nominally significant diseases with sufficient incident cases (defined as ≥ 50 incident events in the test set), we regressed the disease age gap Δ on time to onset using ordinary least squares. A negative slope indicates that participants with a larger gap had a shorter time to subsequent disease onset.

Comparison with existing proteomic clocks and variance decomposition.

We compared MaxIE against the 20 non-sex-specific proteomic clock scores published by Oh et al. [10]: the conventional clock (trained on all proteins), the organismal clock (trained on organ-nonspecific proteins), and 18 organ- and tissue-specific clocks (each trained on proteins annotated to one organ via Genotype-Tissue Expression (GTEx)-derived tissue enrichment criteria[29]). For each disease, the best existing clock was defined as the Oh et al. clock achieving the highest mediation index f^* for that disease; this clock was used consistently across all comparison metrics (f^*, age correlation and AUC) so that the comparison is always against the strongest available competitor. Variance decomposition used sequential R^2 on the held-out test set ($n = 23{,}022$), taking each disease aging clock score as the outcome. The first model regressed the clock score on chronological age alone ($\widehat{M} = \alpha_0 + \alpha_1 A$), giving R^2_{age} that is the coefficient of determination of this model. The second added the 20 non-sex-specific Oh et al. clock scores ($\widehat{M} = \alpha_0 + \alpha_1 A + \sum_{k=1}^{20} \gamma_k C_k$). The incremental $\Delta R^2_{\text{existing}} = R^2_{(2)} - R^2_{(1)}$ measures the additional variance captured by the 20 existing clocks. The remainder $R^2_{\text{unique}} = 1 - R^2_{(2)}$ is unique to the disease aging clock.

Including all 20 existing clocks in the second model gives a conservative upper bound on the variance in the MaxIE score already captured by existing approaches. The head-to-head comparison uses only the single best organ clock per disease.

Weight structure analysis and pathway enrichment.

To characterize the shared versus disease-specific structure of the weight vectors, we computed the 20×20 pairwise Pearson correlation matrix of the weight vectors for the 20 nominally significant diseases. The shared component was defined as the first principal component (PC1) of the $20 \times 2{,}916$ weight matrix. PC1 captures the dominant direction of protein variation shared across all 20 clocks and was used as input to pathway enrichment alongside the individual disease weight vectors. Disease groupings were visualized by hierarchical clustering (average linkage on $1 - r$ distance) and UMAP (n_neighbours=10, min_dist=0.8, cosine metric) using the uwot 0.2-3 R package[30].

Pathway enrichment used pre-ranked gene set enrichment analysis (fgseaMultilevel from the fgsea 1.28 R package) [31] on the absolute weight vectors $|\boldsymbol{w}|$. Absolute values were used because the sign of a MaxIE weight reflects contribution direction to the composite score rather than protein expression direction; testing $|\boldsymbol{w}|$ identifies whether pathway members carry disproportionately large contributions regardless of sign. Gene set databases were obtained from MSigDB v2026.1[32] via the msigdbr 7.5 R package: Hallmark (50 sets), KEGG Legacy (186 sets), Reactome (1,839 sets), and GO Biological Process (7,538 sets). Sets were intersected with the 2,916 Olink proteins; sets with fewer than 10 overlapping proteins were excluded. False discovery rate (FDR) was controlled by the Benjamini–Hochberg procedure[33] within each database independently. Enrichment was run for the shared component (PC1 of the weight matrix) and each of the 20 disease weight vectors.

Linearity of the mediation signal.

Because MaxIE finds the optimal linear mediator $M(\boldsymbol{X}^*) = \boldsymbol{X}^*\boldsymbol{w}$ for any input feature representation $\boldsymbol{X}^*$, nonlinear mediators can be evaluated by constructing a nonlinear $\boldsymbol{X}^* = g(\boldsymbol{X})$ and applying MaxIE in that transformed space. To test whether nonlinear features of the proteins add mediation signal, we applied MaxIE to the augmented matrix $[\boldsymbol{X}, \boldsymbol{X}^*]$ for eleven feature maps in three families.

Extra Trees[34] (500 trees; max_leaf_node=100; min_samples_leaf=10; max_features=$\sqrt{p}$): leaf-node indicators were one-hot encoded, and the resulting sparse matrix passed to the cosine test, trained on randomly permuted outcomes to avoid outcome leakage.

Kernel methods[35]: the method was applied to a radial basis function (RBF) kernel ($\gamma = 1/p$).

Autoencoders[36] and *variational autoencoders (VAEs)* [37]: encoder architectures with two hidden layers (256 and 64 units, ReLU activations) and latent dimension 32, trained unsupervised; latent representations passed to the cosine test.

LightGBM [38]: 200 trees with max_leaf_nodes=50; leaf indicators TruncatedSVD-compressed to 200 dimensions.

For all nonlinear methods, mediation strength was evaluated by computing f^* on the encoded feature matrix. Details and results are provided in Supplementary Note 1 and Supplementary Table 4. Extra Trees, the RBF kernel map and TruncatedSVD used scikit-learn 1.4.

Stanford Alzheimer's Disease Research Center replication cohort.

Independent validation was performed using the Stanford Alzheimer's Disease Research Center (ADRC) Olink Explore 3k cohort[12]. Plasma samples were collected from 501 participants (mean age 70.7 years, s.d. 8.2; age range 36–87 years) with available ADRC data from visit 1 blood draw. Diagnoses were assigned by clinical consensus and comprised healthy controls (HC; $n = 230$),

mild cognitive impairment (MCI; $n = 88$), Alzheimer's disease (AD; $n = 67$), Parkinson's disease (PD; $n = 65$), Lewy body dementia (LBD; $n = 16$), and Parkinson's disease with MCI (PD-MCI; $n = 35$). Plasma proteins were quantified using the Olink Explore 3k platform, yielding 2,909 proteins overlapping with the 2,916 UKB proteins; the 7 non-overlapping proteins were imputed to zero on the standardized scale (equivalent to the UKB training mean), contributing no information to the composite score.

Disease age computation. All weights and calibration parameters were frozen from the UKB training set and applied without modification to the ADRC cohort. Because the UKB and Stanford ADRC cohorts were processed on different Olink platforms and underwent independent preprocessing pipelines, a systematic intercept shift was observed across proteins. Across 2,909 proteins shared between platforms, 80% had lower mean normalized protein expression (NPX) in the ADRC than in the UKB (median z-shift $= -0.10$ UKB s.d. units), and the within-cohort protein s.d. was on average 66% larger in the ADRC (mean 0.985 vs. 0.593 NPX). Direct application of UKB normalization parameters to the ADRC introduced a systematic negative offset of approximately -13 years in the raw disease age gap.

Per-protein intercept batch correction. For each protein j, we fitted the linear model $x_j = a_{0j} + a_{1j}\,A + \varepsilon$ independently in the UKB training set ($n = 21{,}504$) to obtain intercept a_{0j}^{UKB}, and in the ADRC healthy controls ($n = 230$) to obtain a_{0j}^{ADRC}. The batch offset $\mathrm{batch}_j = a_{0j}^{\mathrm{ADRC}} - a_{0j}^{\mathrm{UKB}}$ captures the age-independent platform mean shift for protein j (mean $|\mathrm{batch}| = 0.66$ NPX). For any individual i in the ADRC, the batch-corrected z-score is

$$z_{ij} = \frac{x_{ij} - \mathrm{batch}_j - \mu_j^{\mathrm{UKB}}}{\sigma_j^{\mathrm{UKB}}},$$

where μ_j^{UKB} and σ_j^{UKB} are the frozen UKB training mean and s.d. for protein j. Because batch_j is a fixed per-protein constant estimated once from ADRC healthy controls, it can be applied to any new individual from the same Olink platform without requiring additional cohort-level data. After batch correction, the mean disease age gap for ADRC healthy controls was $+0.98$ years (s.d. 7.6 years), confirming successful centering of the healthy-control distribution without eliminating the disease age signal. For each of the 31 disease aging clocks, the disease age gap was

$$\Delta_i = \frac{M_i - b_0 - b_1 A_i - b_2 S_i}{b_1}, \quad M_i = \boldsymbol{z}_i^{\top} \boldsymbol{w}^{*},$$

where $\boldsymbol{z}_i$ is the batch-corrected z-score vector, $\boldsymbol{w}^{*}$ are the frozen UKB MaxIE weights, and b_0, b_1, b_2 are the frozen UKB calibration coefficients. All downstream group comparisons and correlations use age-adjusted residuals to remove confounding by the age distribution of clinical groups. Amyloid positivity was defined as plasma pTau217 > 0.21 pg/mL and tau positivity as pTau217 > 0.50 pg/mL; neurodegeneration (N+) was defined as neurofilament light chain (NfL) above the 75th percentile of the ADRC cohort. *APOE* genotype was coded as non-carrier ($\varepsilon3/\varepsilon3$, $\varepsilon2/\varepsilon3$, $\varepsilon2/\varepsilon2$), $\varepsilon3/\varepsilon4$, or $\varepsilon4/\varepsilon4$.

Multidomain linear mixed model (LMM) associations. To characterize the biological correlates of the AD age gap, we fitted the same LMM specification outcome ~ age gap + age + sex + (1 | subject) for each of 20 outcomes: 5 plasma biomarkers (phospho-tau 217, phospho-tau 181, $A\beta_{42/40}$ ratio, glial fibrillary acidic protein (GFAP), NfL), 9 cognitive scores (Montreal Cognitive Assessment (MoCA)[39], Craft story immediate and delayed recall, category fluency, digit span forward and backward, Trail Making A and B, Clinical Dementia Rating (CDR) Sum of Boxes[40]), and 6 global structural magnetic resonance imaging (MRI) measures (hippocampal volume, amygdala volume, lateral ventricle volume, white matter hyperintensity volume, total brain

volume, and mean cortical thickness), all intracranial volume (ICV)-normalized where applicable. Biomarker plasma concentrations were log-transformed before z-scoring. All outcomes were z-scored prior to fitting. For participants with a single visit, ordinary least squares were used in place of the LMM. Benjamini–Hochberg FDR correction was applied jointly across all 20 outcomes within each sample group (HC, HC+MCI, HC+MCI+AD) separately. The figure displays nominal p-values for the HC+MCI+AD group.

Prevalence-by-threshold analysis and age matching. For each clock (AD, PD) and each cohort, we computed disease prevalence as a function of an age gap threshold x: at each x, the prevalence is the fraction of subjects with age gap $\geq x$ who are cases, with 95% binomial confidence intervals computed by the exact Clopper–Pearson method. Thresholds were evaluated at the 5th to 90th percentiles of the pooled age gap distribution in 2% increments (requiring ≥ 15 subjects above each threshold). To enable direct comparison with the UKB test set despite a ≈14-year difference in mean age and a fourfold difference in case prevalence, UKB subjects were stratified-downsampled to match the ADRC age distribution. For each outcome (AD clock, PD clock), the UKB test set was restricted to the healthy control group and the corresponding case group. Subjects were binned into 2-year age intervals, applied separately to controls and cases. Within each (age bin × diagnostic group) cell, we retained $\min(n_{\mathrm{UKB}}, n_{\mathrm{ADRC}})$ randomly selected UKB subjects (random seed fixed to 42 for reproducibility). This joint downsampling simultaneously matches the within-group age distribution and naturally reproduces the ADRC HC:case ratio without a separate ratio-adjustment step. All prevalence curves are computed on this matched UKB sample.

Regional brain imaging analysis. Cortical parcellation used the Desikan–Killiany (DK) atlas [41], which assigns each hemisphere's cortical surface to 34 named regions based on FreeSurfer surface reconstructions[42]. For each region, the left- and right-hemisphere measurements were averaged to obtain a bilateral value, halving the number of statistical comparisons and reducing hemisphere-specific noise. Three imaging modalities were analyzed: amyloid positron emission tomography (PET; mean cortical standardized uptake value ratio, SUVR), tau PET (SUVR), and structural MRI-derived cortical thickness.

Visit matching. Olink proteomic visit records were linked to imaging records in two steps. First, each Olink visit was matched to the ADRC diagnosis record with the nearest chronological age within the same subject (no maximum age gap threshold applied), yielding a paired diagnosis-visit index. Second, the imaging scan closest in visit number to this matched diagnosis visit was selected within subject, minimizing temporal mismatch between proteomic and imaging measurements.

Analysis sub-samples. Non-AD diagnoses (Parkinson's disease, Lewy body dementia, vascular dementia, and frontotemporal dementia) were excluded from all imaging sub-samples. Two sub-samples were defined: (i) an early-stage HC+MCI sub-sample restricted to healthy controls and subjects with mild cognitive impairment ($n_{\mathrm{obs}} = 88$, $n_{\mathrm{subj}} = 79$), and (ii) a broader HC+MCI+AD sub-sample that additionally includes subjects with confirmed Alzheimer's disease dementia ($n_{\mathrm{obs}} = 100, n_{\mathrm{subj}} = 91$).

Statistical model. For each of the 34 bilateral cortical regions and each imaging modality, we fitted the linear mixed model

$$y_{ij}^{(\mathrm{std})} = \beta_0 + \beta_1 \, \Delta\hat{a}_i^{(\mathrm{std})} + \beta_2 \, a_i^{(\mathrm{std})} + \beta_3 \, \mathrm{sex}_i + u_i + \varepsilon_{ij},$$

where $y_{ij}^{(\mathrm{std})}$ is the standardized bilateral regional measure for subject i at visit j; $\Delta\hat{a}_i^{(\mathrm{std})}$ is the standardized AD disease age gap (residualized for chronological age within the ADRC); $a_i^{(\mathrm{std})}$ is standardized chronological age; sex_i is binary (F/M); and $u_i \sim \mathcal{N}(0, \sigma_u^2)$ is a subject-level random intercept. All continuous predictors and outcomes were z-scored before fitting, so that β_1

is a standardized effect size directly comparable across regions and modalities. Models were fitted by maximum likelihood (REML = FALSE) using lme4 1.1-35 [43] and lmerTest 3.1-3[44] with the bobyqa optimizer (maximum iterations: 2×10^5); p-values for the age gap coefficient use Satterthwaite approximate degrees of freedom. For subjects with only a single observation, ordinary least-squares regression replaced the LMM. Regions with fewer than 15 valid observations were excluded. Multiple-testing correction used the Benjamini–Hochberg FDR procedure within each imaging modality independently. For cortical thickness, $\beta_1 < 0$ indicates thinning associated with higher age gap; brain surface visualizations use sign-negated colors so that red indicates thinning, following neuroimaging convention. Brain surface maps were rendered with ggseg 1.6 [45] (DK atlas).

Global Neurodegeneration Proteomics Consortium SomaScan replication.

Plasma SomaScan v4.1 data[46] were obtained from the Global Neurodegeneration Proteomics Consortium (GNPC)[23], comprising 17,269 samples from 13,449 participants across 17 contributing sites. We retained the first visit per participant with a valid recorded age, yielding 13,388 individuals, of whom 13,388 had a six-group diagnostic label. Proteomic preprocessing and diagnosis harmonization followed the consortium pipeline and are detailed in the Supplementary Note 2. We restricted the analysis to the 2,684 SomaScan aptamers whose target proteins are represented on the UKB Olink panel, matched by UniProt accession[47], providing a comparable feature space for cross-platform replication. Aptamer values were harmonized across sites with ComBat[48], using site as the batch and preserving age and sex while excluding diagnosis[49], and were then standardized. Cohort characteristics, diagnostic composition and per-site structure are provided in Supplementary Table 8.

An AD aging clock was trained in GNPC. Clock fitting used AD cases and cognitively normal controls only, and no diagnostic label entered ComBat, so the out-of-fold gaps of participants with mild cognitive impairment, Parkinson's disease, frontotemporal dementia or amyotrophic lateral sclerosis are read-outs on groups the model never saw. Cross-fitting was fivefold and grouped by site (GroupKFold): within each fold the MaxIE weights and the calibration of the composite score on chronological age and sex were estimated in the training sites and applied to the held-out sites, so no participant's own site contributed to the weights or the calibration applied to them. ComBat harmonization was applied once to the full cohort and was not fold-specific. The cosine test was evaluated in the 8,683 AD cases and controls rather than in all participants, because the GNPC clock is defined on that contrast. Out-of-fold gaps were summarized across the six diagnostic groups, and site robustness was assessed by recomputing the pooled AD-minus-control difference with one site omitted at a time, across the ten sites contributing at least 10 AD cases and 10 controls.

Acknowledgements

Stanford ADRC data collection was supported by the Stanford Alzheimer's Disease Research Center, NIH/NIA grant P30 AG066515. Z.H. was supported by NIH/NIA awards AG089509, AG066206, and AG066515. J.T. and R.M.W. were supported by the National Library of Medicine (NLM) T15 Training Grant (T15LM007033).

Author contributions

Z.H. and J.P. designed and performed the analyses, generated the figures, and wrote the manuscript. J.T., R.M.W. and J.F.W. verified the analyses. K.Y., Y.L.G., E.C.M., V.W.H., and T.W.-C. provided critical review of the manuscript. Z.H. conceived the study and supervised all the analysis.

Competing interests

The authors declare no competing interests.

Data availability

UK Biobank data are available to approved researchers through the UK Biobank Access Management System (https://www.ukbiobank.ac.uk) under application number 45420. The Stanford ADRC data used in this study are available upon reasonable request to the Stanford ADRC Data Release Committee (https://web.stanford.edu/group/adrc/cgi-bin/web-proj/datareq.php). GNPC data are available upon request following the standard protocol (www.neuroproteome.org/harmonized-data-set-hds).

Code availability.

The codes used for analyses in the present study are available at the following link: https://github.com/Junkkkk/disease_age_clock. The optmed package is available at https://github.com/statzihuai/optmed.

Figures

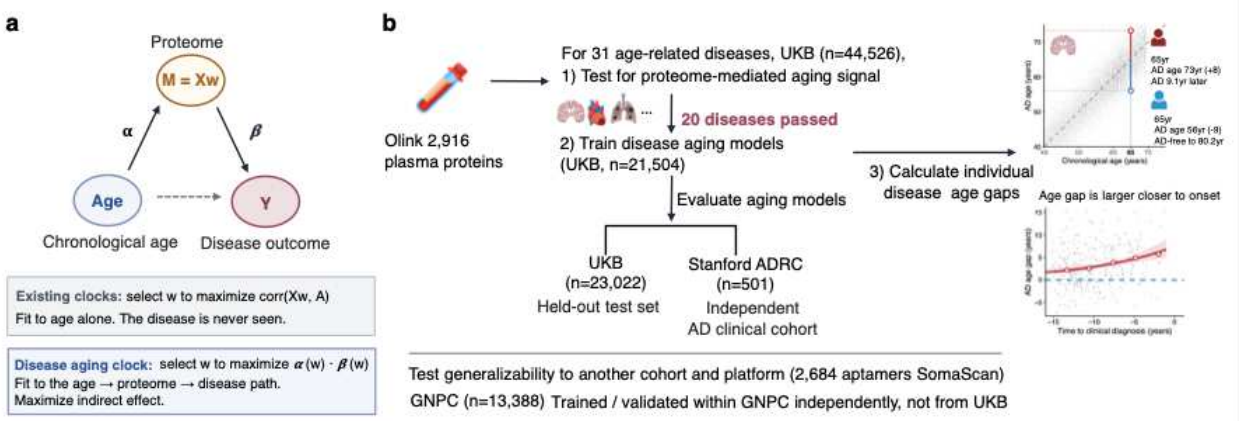
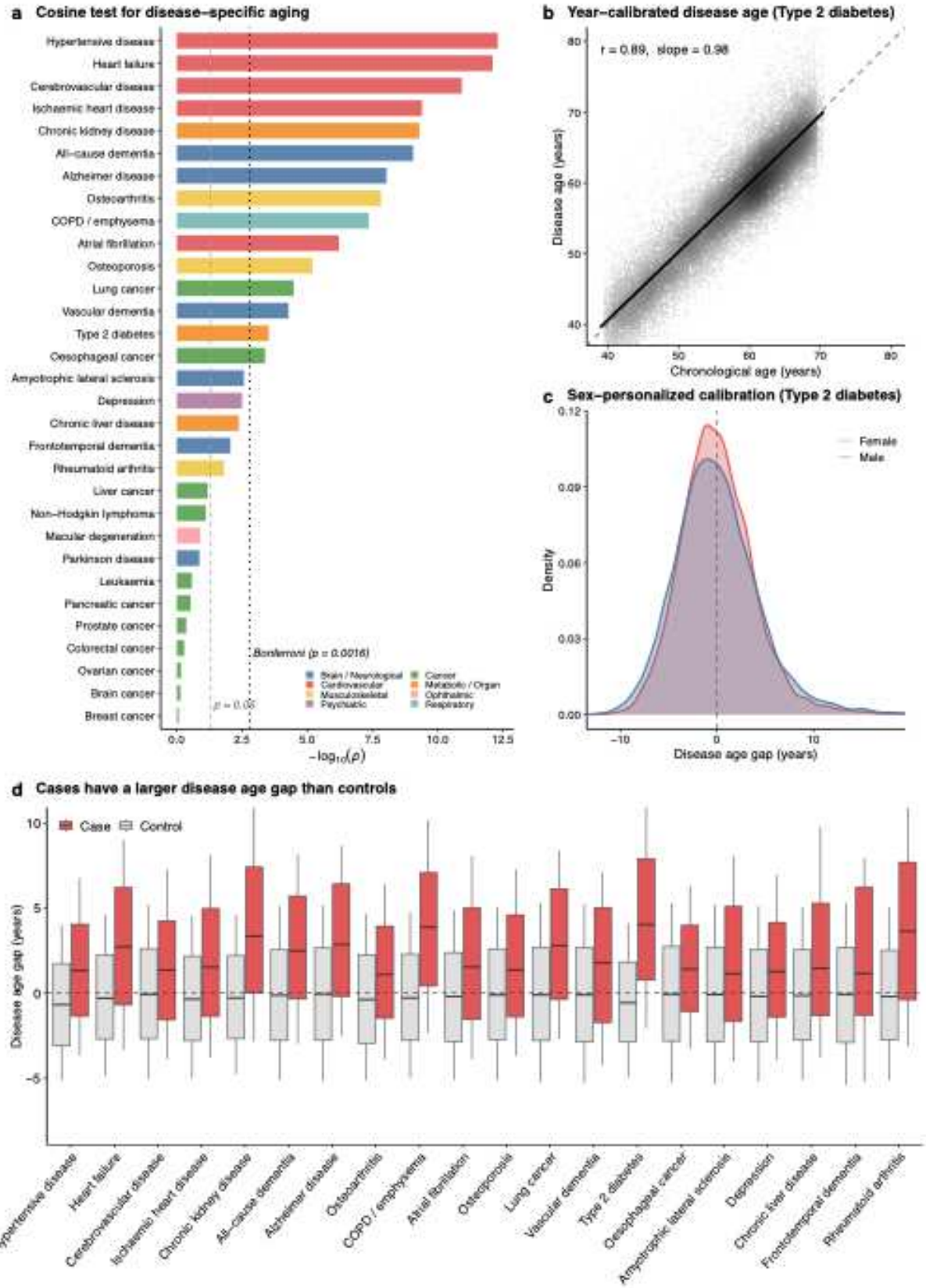


Figure 1. Overview of the disease aging clock framework.

a) Conceptual framework. A weighted plasma-protein composite is defined as M(w)=Xw, where X is the matrix of measured protein abundances and w is a vector of protein weights. Chronological age (A) is associated with disease outcome (Y) both directly and indirectly through M(w). The coefficient α(w) represents the association between chronological age and the protein composite, whereas β(w) represents the association between the protein composite and disease outcome conditional on chronological age. Conventional aging clocks select w to maximize the correlation between M(w) and chronological age. The disease aging framework instead selects w to maximize the age-mediated disease signal, quantified by $\alpha(w)\beta(w)$.

b) Study design for disease age model development and evaluation. Plasma levels of 2,916 proteins were measured in 44,526 UK Biobank (UKB) participants. Thirty-one age-related diseases were first tested for evidence of a proteome-mediated aging signal using the cosine test, and 20 diseases were carried forward for disease age model development. Participants were divided into geographically defined training (n = 21,504) and held-out test (n = 23,022) sets. Disease-specific protein weights were estimated in the training set and applied to the held-out test set to estimate disease age and the disease age gap. For Alzheimer's disease (AD), the UKB-derived model was directly applied to an independent Stanford Alzheimer's Disease Research Center (ADRC) cohort (n = 501). To evaluate cross-platform reproducibility, an independent AD disease age model was trained and evaluated in the Global Neurodegeneration Proteomics Consortium (GNPC; n = 13,388) using 2,684 SomaScan aptamers.



867
868 **Figure 2. Construction and validation of disease aging models.**

a) Cosine test results for 31 age-related diseases in 44,526 UKB participants. Bars show $-\log_{10} P$ values and are colored by disease category. The dashed line indicates nominal significance ($p = 0.05$) and the dotted line indicates the Bonferroni-corrected significance threshold ($p < 0.0016$). Fifteen diseases passed Bonferroni correction and 20 reached nominal significance and were carried forward for disease age model construction.
b) Year-calibrated disease age plotted against chronological age for the type 2 diabetes (T2D) model in the held-out test set (n = 23,022). Dashed red line shows the identity (y = x); solid blue line shows the regression fit (r = 0.89, slope = 0.98).
c) Distribution of the T2D age gap after sex-personalized calibration. The disease age gap was defined as disease age minus chronological age. Distributions for females and males were centered at zero; the overall standard deviation was 4.1 years.
d) Disease age gap in prevalent cases (red) and controls (gray) for the 20 diseases that reached nominal significance.

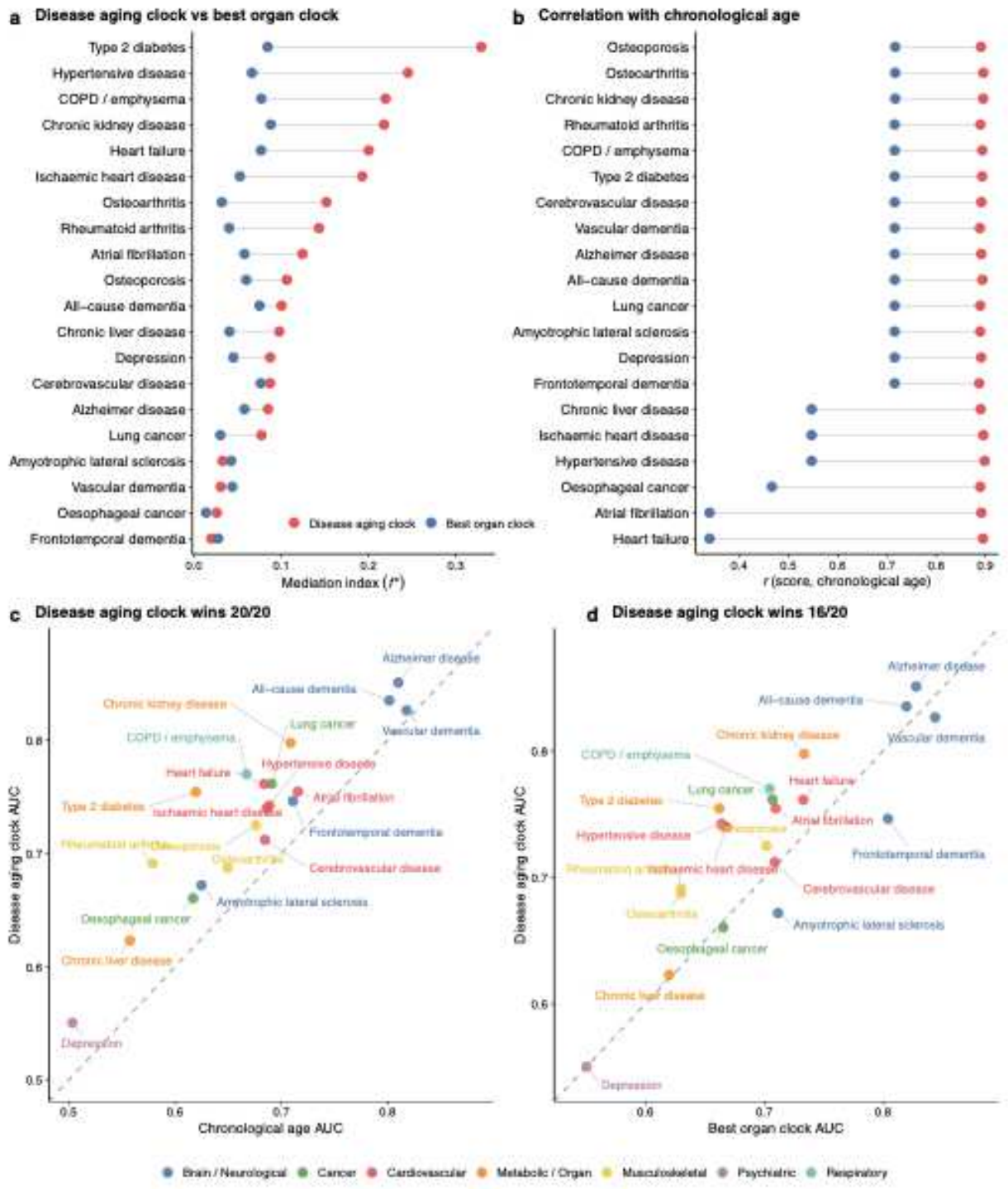


Figure 3. Disease aging clocks outperform organ-specific clocks across disease-relevant metrics. Restricted to the 20 nominally significant diseases (Supplementary Fig. 1 shows all 31). For each disease, the best-performing organ-specific clock was selected from the organ aging models of Oh et al.[10] based on the highest mediation index (f^*) and used consistently across all comparison metrics.

a) Comparison of the mediation index (f^*) between the disease aging clock (red) and the best organ-specific clock (blue) for each disease.

b) Correlation of each clock with chronological age (r=0.89–0.90 for disease aging clocks and r=0.34–0.72 for the organ clock with the highest f^* for each disease).
c) Disease classification performance (AUC) of disease aging clocks compared with chronological age alone as a classifier.
d) Disease classification performance (AUC) of disease aging clocks compared with the best organ-specific clock. Variance decomposition is shown in Extended Data Fig. 1.

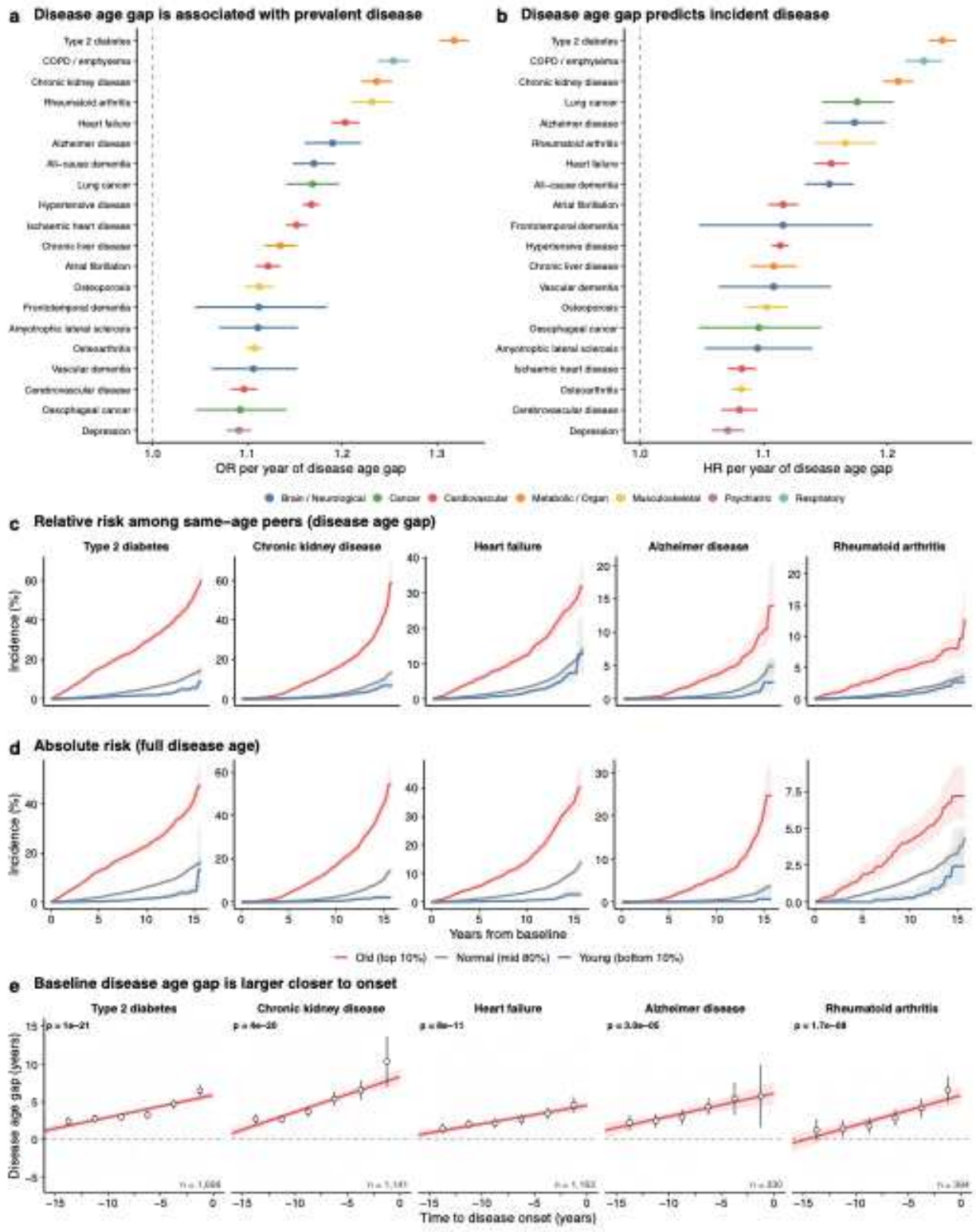


Figure 4. Disease age gap predicts future disease onset and tracks preclinical progression. **a)** Cross-sectional odds ratio (OR) per year of disease age gap for the 20 nominally significant diseases, adjusted for chronological age in the held-out UKB test set (n=23,022).

b) Prospective hazard ratio (HR) per year of disease age gap for incident disease estimated using Cox proportional hazards models adjusted for chronological age (median follow-up 10.4 years). **c)** Kaplan–Meier cumulative incidence curves stratified by disease age gap for five representative diseases (Young: bottom 10%; Normal: middle 80%; Old: top 10%). Because the age gap is independent of chronological age after calibration, differences in cumulative incidence reflect relative disease risk among individuals of the same chronological age. Shaded bands indicate 95% confidence intervals. Supplementary Fig. 3 shows all 20 diseases. **d)** Kaplan–Meier cumulative incidence curves stratified by disease age using the same percentile thresholds as in c. Because disease age incorporates chronological age, cumulative incidence reflects absolute disease risk. Supplementary Fig. 4 shows all 20 diseases. **e)** Baseline disease age gap versus time to disease onset in future incident cases, for five representative diseases. Grey points are individual cases. Open circles are means within 2.5-year bins with 95% CI error bars, shown out to 15 years before onset for bins. Red lines are linear regressions of age gap on time to onset fit to all individual cases. The p values refer to these fits. The y axis spans −6 to 15 years, excluding 1.2–3.8% of cases below and 0.9–4.5% above depending on disease. All 20 diseases in Extended Data Fig. 2.

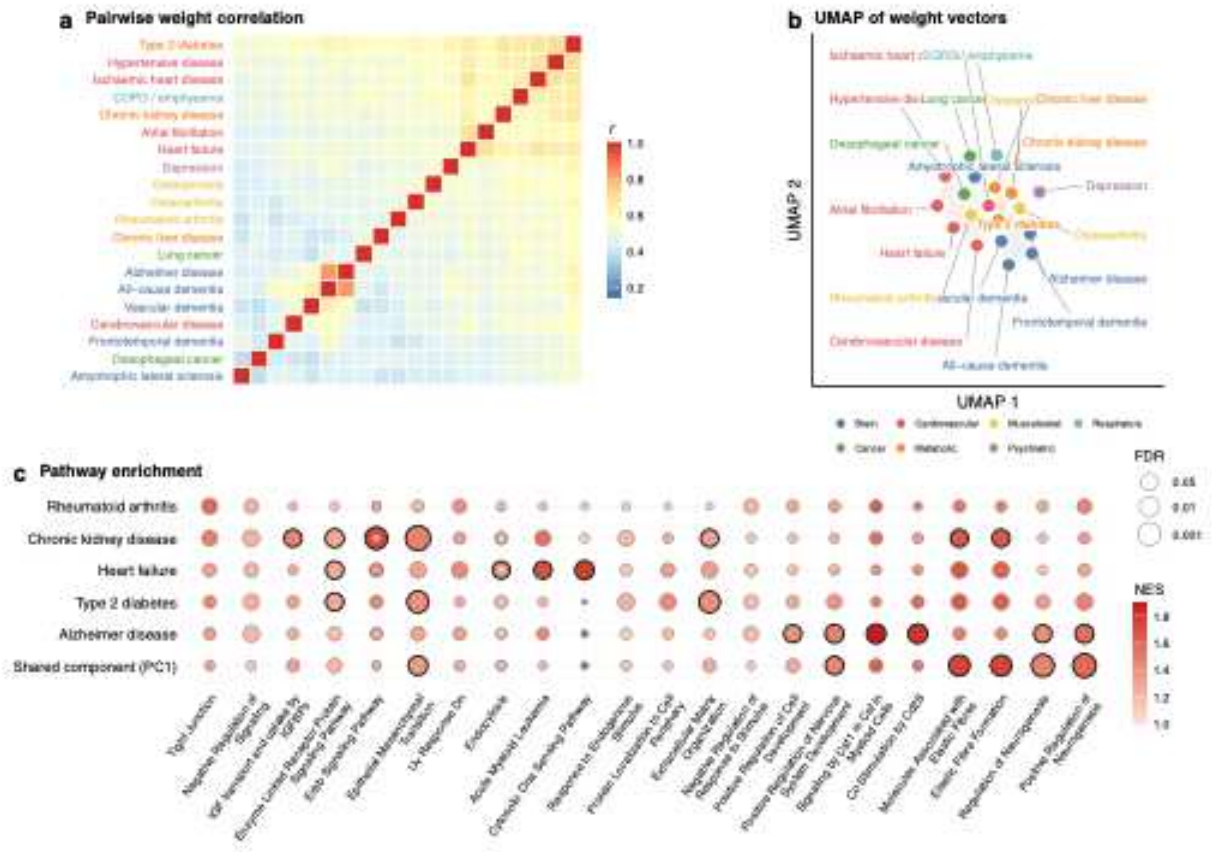


Figure 5. Disease aging clock weights reveal shared and disease-specific aging biology.
a) Pairwise Pearson correlations among disease aging clock weight vectors for the 20 diseases. Weight vectors were hierarchically clustered, and disease labels are colored by disease category.
b) Uniform manifold approximation and projection (UMAP) of disease aging clock weight vectors in the 2,916-dimensional protein space using cosine distance. Each point represents one disease and is colored by disease category.
c) Pathway enrichment analysis for the shared component (first principal component; PC1) and five disease aging clocks using fgsea performed on absolute protein weights. Dot size represents $-\log_{10} FDR$, fill color indicates the normalized enrichment score (NES) and black outlines denote false discovery rate (FDR) < 0.05. Full results are provided in Supplementary Table 3.

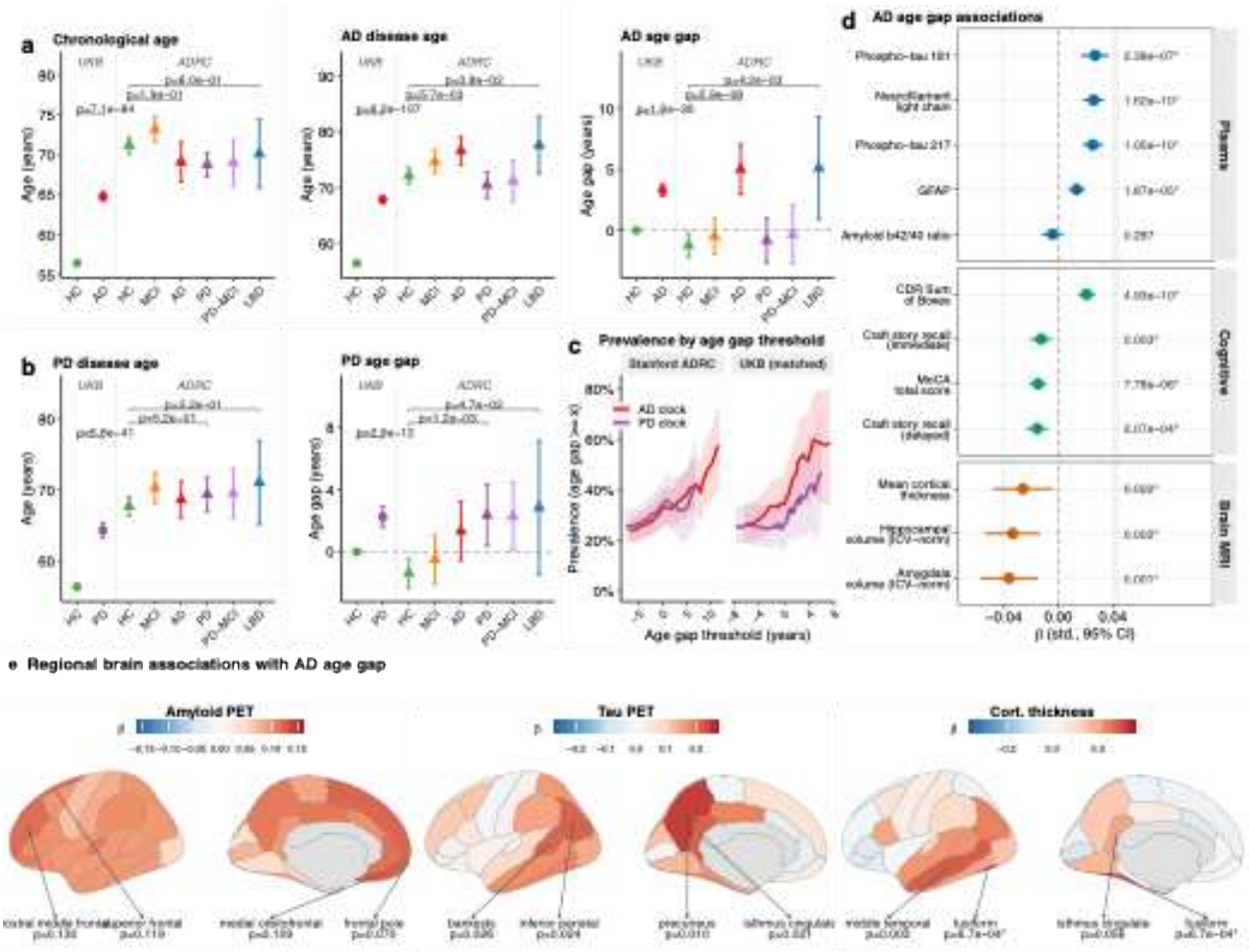


Figure 6. External validation of the AD aging clock in the Stanford Alzheimer's Disease Research Center(ADRC).

a) Cross-cohort calibration of the UK Biobank (UKB)-derived AD aging clock. Mean chronological age, disease age and disease age gap are shown across diagnostic groups in the UKB test set (circles; n=23,022) and the independent Stanford ADRC cohort (triangles; n=501). Points and error bars indicate means and 95% confidence intervals. Brackets indicate two-sided Wilcoxon rank-sum tests.

b) Specificity analysis using the UKB-trained Parkinson's disease (PD) aging clock. Mean PD disease age and the disease age gap are shown across the same diagnostic groups.

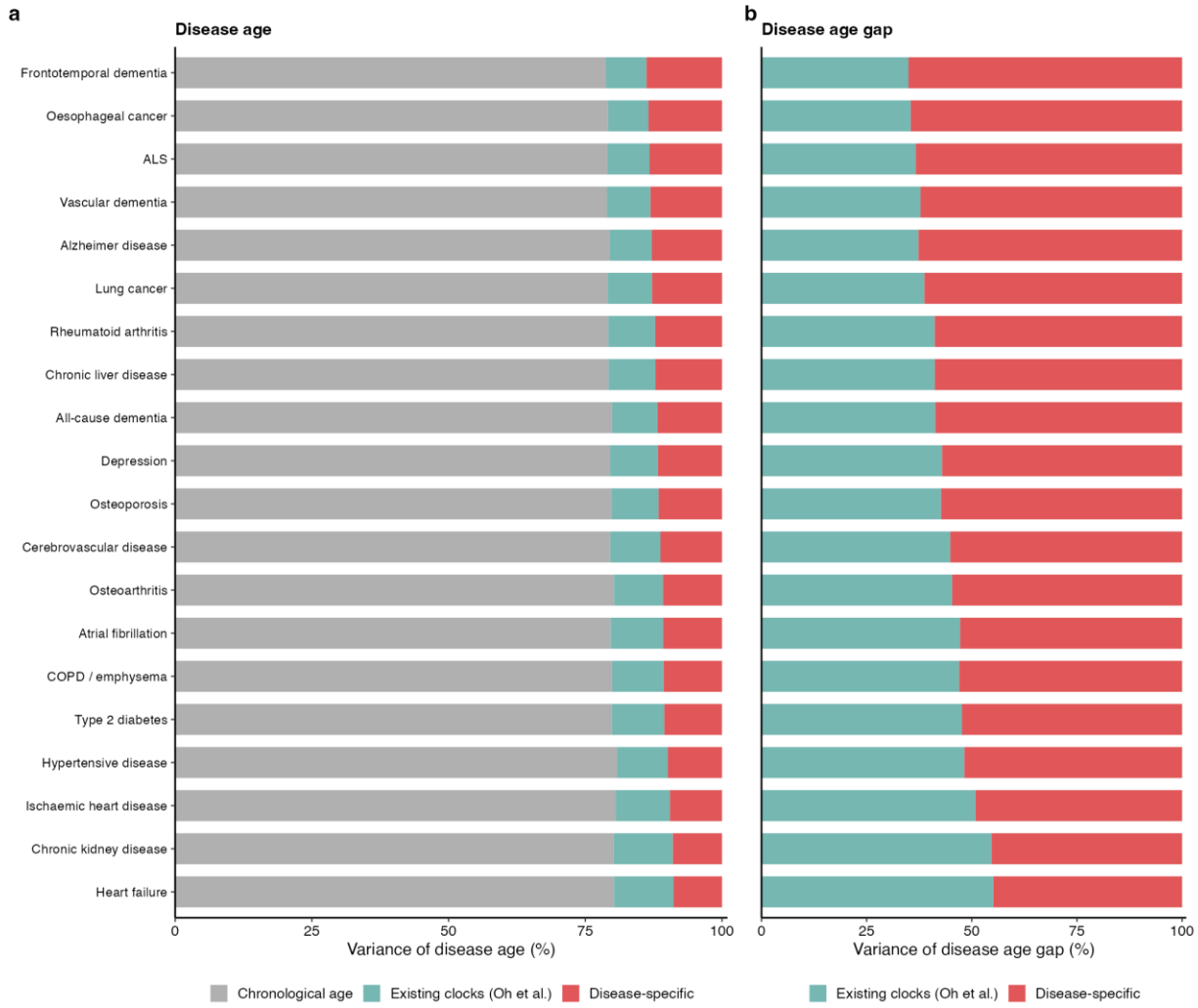
c) Disease prevalence among participants whose disease age gap exceeds threshold x, shown for the AD and PD aging clocks in the Stanford ADRC and an age-matched subset of the UKB test set. Shaded bands indicate 95% exact binomial confidence intervals (CI).

d) Standardized regression coefficients and 95% CIs for associations between the AD disease age gap and plasma biomarkers, cognitive measures and structural MRI phenotypes in the full diagnostic sample comprising healthy controls (HC), mild cognitive impairment (MCI) and AD. Linear mixed-effects models included disease age gap, chronological age and sex as fixed effects and participant as a random effect.

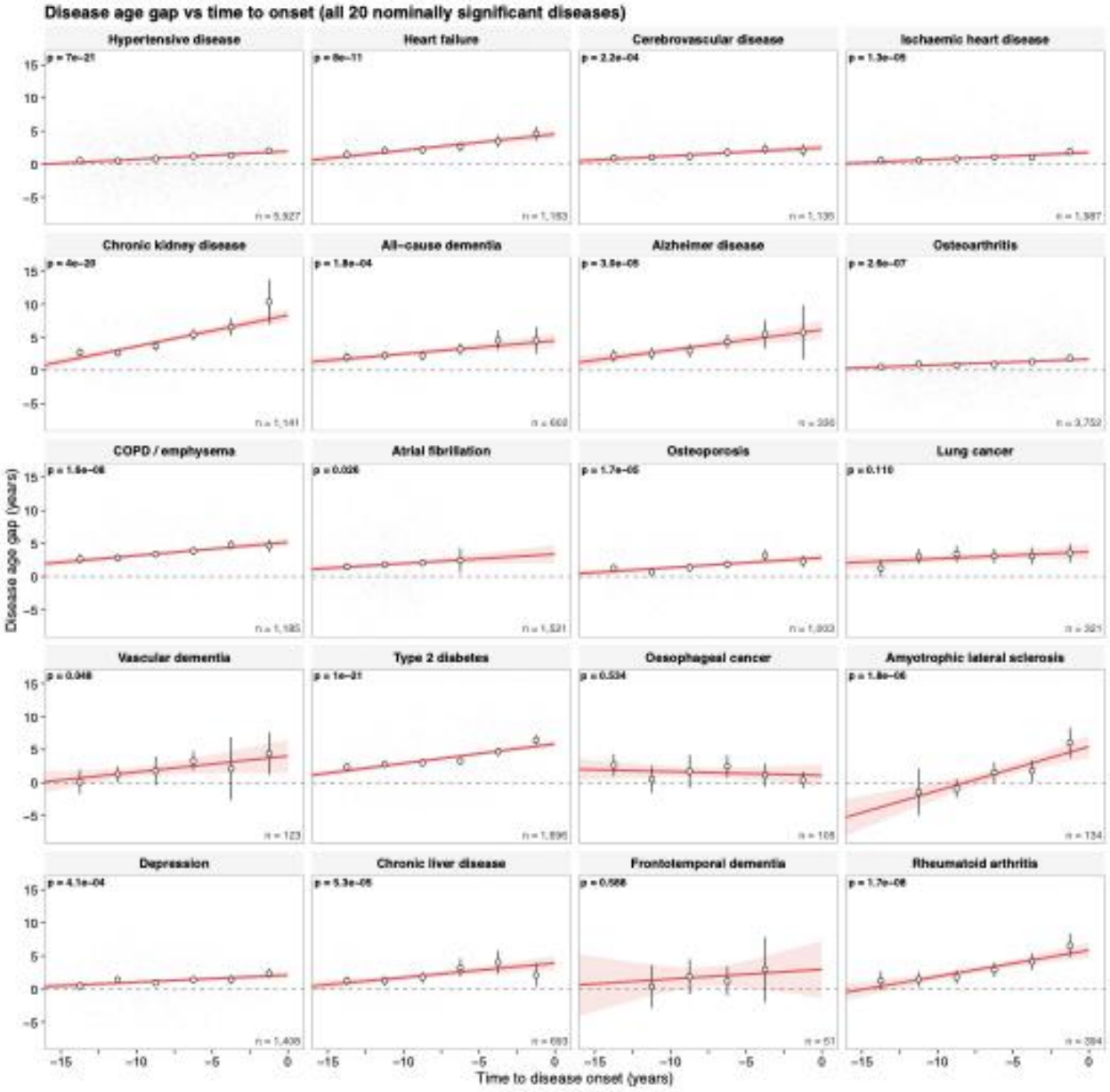
e) Regional associations between the AD disease age gap and amyloid PET, tau PET and cortical thickness in the full diagnostic sample comprising HC, MCI and AD, mapped to the Desikan–Killiany atlas. Regression coefficients are standardized. Cortical-thickness coefficients are sign-

954 inverted so that positive values indicate greater cortical thinning. The two strongest associations
955 in each view are labeled. * indicates FDR $q < 0.05$.

Extended Data Figures

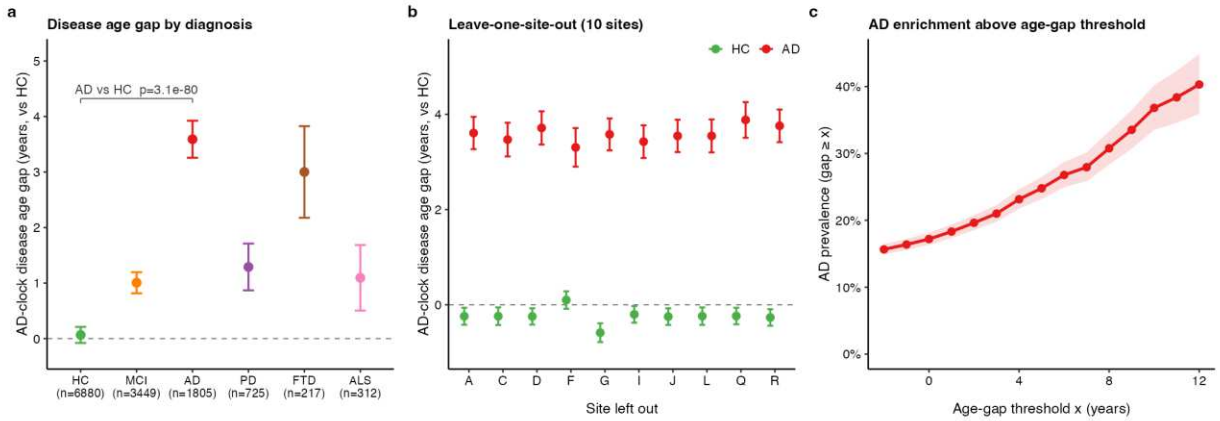


Extended Data Fig. 1. Variance decomposition of disease aging clocks. Both panels use sequential R^2 in the held-out test set (n = 23,022).
a) Decomposition of disease age. Stacked bars show the variance explained by chronological age alone (grey; mean 79.7%, range 78.8–80.9% across the 20 diseases), the additional variance explained by the 20 non-sex-specific proteomic aging signatures of Oh et al.[10] (teal; mean 8.8%, range 7.4–10.9%), and the residual variance unique to each disease aging clock (red; mean 11.5%, range 8.9–13.8%).
b) Decomposition of the disease age gap, which is orthogonal to chronological age by construction. The gap accounts for a mean of 20.3% of the variance in disease age (range 19.2–21.3%); of that, a mean of 43.6% is captured by those 20 signatures (teal) and 56.4% is residual variance not linearly explained by them (red).



Extended Data Fig. 2. Baseline disease age gap versus time to disease onset for all 20 nominally significant diseases.

Each point represents one future incident case (24,671 cases in total), open circles the mean disease age gap within 2.5-year bins with 95% confidence intervals. The x axis shows the time from the proteomic blood draw to clinical disease onset, reversed so that onset approaches from left to right; the y axis shows the disease age gap measured at baseline. Red lines show linear regression fits with 95% confidence intervals. A positive slope indicates that participants whose disease onset was more imminent had larger disease age gaps at baseline, consistent with the gap reflecting the stage of a progressing disease process. The slope is positive for 19 of 20 diseases and reaches $P < 0.05$ for 17. Representative diseases are shown in Fig. 4e.



Extended Data Fig. 3. Cross-platform replication of the AD aging clock in GNPC SomaScan data.

a) Disease age gap by diagnostic group (mean ± 95% CI), expressed relative to the control median: HC (n = 6,880), MCI (n = 3,449), AD (n = 1,805), PD (n = 725), frontotemporal dementia (n = 217) and amyotrophic lateral sclerosis (n = 312). The AD mean gap was 3.59 years (95% CI 3.26–3.92), a median difference of 3.31 years from HC (two-sided Wilcoxon rank-sum $P = 3.1 \times 10^{-80}$). **b)** Leave-one-site-out robustness of the replication. Each x position shows the mean disease age gap (± 95% CI) for controls (green) and AD cases (red) recomputed after omitting the labelled site, across the ten sites contributing at least 10 of each group (n = 7,005; 1,799 cases and 5,206 controls). The pooled AD control difference was 3.84 years (95% CI 3.46–4.21) and ranged from 3.21 years (omitting site F) to 4.16 years (omitting site G), so no individual site drives the replication. **c)** Fraction of AD cases among participants with a disease age gap at or above a given threshold, with a 95% Wilson confidence band. Prevalence rises from 15.6% at a threshold of −2 years to 40.3% at 12 years.

Supplementary Files

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

- supplementarytables261106.xlsx
- supplementarynote261106.docx