Distance of covariance (DISCO), a novel measure of network homeostatic dysregulation, reveals organ system interconnections underlying mortality and disease risk

Meng Hao[1,2,3#*], Hui Zhang[2,4#], Yi Li[1,2#], Yaqi Huang[2], Jingyi Wu[2], Shuishan Zhang[2], Zixin Hu[2], Xiangnan Li[2], Shuai Jiang[5], Kamaryn T. Tanner[6], Jie Chen[1], Zhijun Bao[1], Jiucun Wang[2*], Alan A Cohen[6,7]*, Yiqin Huang[1*], Li Jin[2], Xiaofeng Wang[2*].

[1] Shanghai Key Laboratory of Clinical Geriatric Medicine, Shanghai Institute of Geriatrics and Gerontology, Huadong Hospital, Shanghai Medical College, Fudan University, Shanghai, 200040, China

[2] Human Phenome Institute, Zhangjiang Fudan International Innovation Centre, Fudan University, Shanghai, 201203, China

[3] Fudan Zhangjiang Institute, Shanghai 201203, China

[4] School of Global Health, Chinese Center for Tropical Diseases Research, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, China

[5] Department of Vascular Surgery, Shanghai Key Laboratory of Vascular Lesion Regulation and Remodeling, Shanghai Pudong Hospital, Fudan University Pudong Medical Center, Shanghai, 201399, China.

[6] Butler Columbia Aging Center, Mailman School of Public Health, Columbia University, NY 10032, America

[7] Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, NY 10032, America

# These authors contributed equally.

* Corresponding author:

Meng Hao, Email: haombio@gmail.com

Xiaofeng Wang, E-mail: wangxiaofeng@fudan.edu.cn.

Yiqin Huang, yiqin_huang@126.com

Alan A Cohen, E-mail: aac2277@cumc.columbia.edu

Jiucun Wang, E-mail: jcwang@fudan.edu.cn

Abstract

Aging manifests as the progressive declines of homeostatic resilience and repair mechanisms, marked by dysregulations across systems and increasing individual heterogeneity. However, the breadth of measures of homeostatic dysregulation remains underexplored. Here, we introduce DISCO as a novel measure of homeostatic dysregulation, integrating clinical, proteomics, metabolomics, and microbiomes data. DISCO demonstrated moderate correlation with chronological age but robustly predicted mortality, frailty, and chronic disease risk, outperforming Mahalanobis distance in health outcome prediction, comparable to the best epigenetic clocks. Organ/tissue-specific DISCO analysis revealed limited organ-disease specificity, suggesting systemic rather than localized dysregulation drives health decline. Network analysis identified aging-associated proteins as central hubs strongly linked to DISCO scores; further, organ-level DISCO metrics most predictive of age and outcomes were also central within biological networks. Collectively, DISCO emerges as a validated measure of whole-body homeostatic dysregulation, providing a tool for aging risk stratification and insights into systemic aging mechanisms.

Introduction

Recently, there is renewed interest in the old hypothesis that aging is the result of an increase in entropy in biological systems, and of the increasing inability of homeostatic and repair mechanisms to resist this increase in disorder[1-4]. This model predicts that numerous upstream causes initiate positive feedback loops which traverse the boundaries of classically defined biological systems, with impacts at diverse scales (molecular, cellular, tissue, organ, integrative) as well as across specific classes of cells, tissues, and organs. Under this model, increasing entropy should result in dysregulation of various biological systems, manifested as increasing heterogeneity. While dysregulation of specific systems and organs may be measurable, such dysregulations are likely to have impacts across systems due to impaired cross-talk and coordination.

Increasing heterogeneity and dysregulation with age have indeed been widely observed[5-10]. The best-known metric of organism-level homeostatic dysregulation is Mahalanobis distance (D_M), which quantifies how far off an individual's profile is from the population mean[11,12]. D_M has largely been applied to widely available data on clinical labs (blood cell counts and biochemistries) and has been shown to increase with age and predict mortality, frailty, and incidence of many chronic diseases, as well as to be exacerbated by various lifestyle and environmental exposures and to mediate the effects of such exposures on health outcomes[9,10,12-15]. However, the breadth of the applicability of measures of homeostatic dysregulation is not well characterized.

More recently, beyond D_M, network metrics demonstrate significant potential in quantifying homeostatic dysregulation by capturing structural perturbations within biological systems[16]. Extensive environmental exposures may profoundly disrupt biological markers and their interconnections[17], suggesting altered correlation patterns could serve as indicators of systemic dysregulation. Furthermore, the topological features of these networks have revealed strong

links to critical health outcomes including functional decline, frailty, disability onset, and mortality risk[18,19]. This underscores the promise of biomarker correlation patterns as potential measures of homeostatic dysregulation.

Here, we introduce a novel approach to quantifying homeostatic dysregulation, the DIStance of COvariance (DISCO, **Fig. 1a**). DISCO works by calculating the correlation matrix for a set of biological markers with and without the inclusion of a target individual whose dysregulation is to be estimated. If the individual has a relatively average profile, these two matrices should be nearly identical. The more unusual the profile, the more the two matrices differ. The distance between the matrices is thus a proxy for homeostatic dysregulation. We apply DISCO in multiple datasets, to multiple types of data, and to predict multiple health outcomes, comparing it to both D_M and to epigenetic clocks (**Fig. 1b**). Furthermore, we generate organ- and tissue-specific versions of DISCO and test their specificity for the prediction of chronic disease incidence and cause-specific mortality from related systems. Finally, we assess networks of dysregulation across systems.

Results

DISCO of clinical biomarkers reflects physiological homeostatic dysregulation

We hypothesized that greater deviations in phenotype correlation (i.e., covariance) from a healthy reference population could indicate more severe disruptions in physiological homeostasis (**Fig. 1a**). Using single-sample analysis, we were able to assess the dysregulation of homeostasis in individual samples by observing how the inclusion or exclusion of a specific sample affected the overall correlation structure. Consequently, we introduced the Distance of Covariance (DISCO) as a differential measure for single samples to quantify the disparity between the perturbed and reference covariance matrices, and applied it across various cohorts and data types (**Fig. 1b**).

In NHANES, DISCO analysis included 27,059 participants (mean age: 49.46 ± 18.02 years) with 24 clinical biomarkers, while UK Biobank (UKB) included 487,517 participants (mean age: 56.54 ± 8.09 years) with 58 biomarkers. Young adults (NHANES: n = 4,734, aged 20–29; UKB: n = 5,282, aged 37–40) served as the reference group to establish biomarker covariance baselines. After log-transformation, DISCO values followed near-normal distributions (NHANES: 0.29 ± 1.54; UKB: -0.24 ± 1.14) and correlated positively with chronological age (**Fig. 1c**, NHANES Spearman $\rho = 0.53$; **Fig. 1d,** UKB $\rho = 0.27$) and strongly with Mahalanobis distance (DM; NHANES $\rho = 0.91$; UKB $\rho = 0.8$). Elevated DISCO levels predicted higher mortality risk after adjustment for age and sex (adjusted hazard ratio (aHR) per unit, NHANES: 1.71, 95% CI: 1.68–1.75; UKB aHR: 1.61, 95% CI: 1.60–1.62). When using age as the time scale DISCO and DM low risk and high risk groups had similar survival probabilities (**Fig. 1e-f**). DISCO demonstrated mortality prediction performance comparable to DM in both NHANES and UKB cohorts, achieving similar concordance indices (C-index: DISCO = 0.767 vs. DM = 0.766 in NHANES; DISCO = 0.670 vs. DM = 0.655 in UKB).

For mortality prediction benchmarking (**Fig. 1g-h, Supp fig. 1**), DISCO was evaluated against established epigenetic clocks (e.g., HorvathAge, GrimAge) in the NHANES cohort, which included 2,532 participants (mean age 66.13 ± 10.08 years) with a median follow-up of 17.2 years (IQR: 9.3–18.8). When combined with chronological age, DISCO exhibited predictive accuracy approaching that of GrimAge (C-index: Age+DISCO = 0.751 vs. GrimAge = 0.759; HorvathAge = 0.713) for mortality. Age-adjusted residual analysis ("Accel") further validated DISCO's predictive capability independent of age, ranking second only to GrimAge and its derivatives. Kaplan-Meier analysis across age groups revealed that DISCO-defined high-risk and low-risk stratification produced the most pronounced survival differences compared to risk groups defined by other epigenetic clocks.

Next, we applied DISCO in three validation cohorts: CHARLS, RuLAS, and CLHLS (**Extended Fig. 1**). At baseline, 11,601 participants (59.46 ± 9.57 years) in CHARLS, 1821 participants (77.02 ± 4.27 years) in RuLAS, and 2213 participants (85.58 ± 12.06 years) in CLHLS were analyzed. The DISCO analysis of routine clinical biomarkers revealed that DISCO levels increased with age across all three cohorts. Additionally, we found a strong correlation between DISCO and Mahalanobis Distance (DM), further indicating DISCO's effectiveness in reflecting dysregulation of homeostasis. In each cohort, elevated DISCO levels were significantly associated with mortality risks after adjustment for age and sex: CHARLS (aHR per unit of DISCO: 1.26, 95% CI: 1.21-1.30), RuLAS (aHR: 1.33, 95% CI: 1.21-1.46), and CLHLS (aHR: 1.27, 95% CI: 1.21-1.33). Kaplan-Meier curves showed that the participants in the lowest and higher quartiles of DISCO faced greater different survival probability across age, which was similar to those categorized via DM across all cohorts.

Refining physiological homeostasis with multi-omics biomarkers

We further utilized DISCO on blood serum proteomics data to determine its ability to reflect physiological homeostasis by integrating a substantial number of proteins from the UKB. Following quality control and feature selection, DISCO analysis included 47,859 participants with 367 proteins. The proteomics DISCO demonstrated a positive correlation with age (ρ = 0.42, **Fig. 2a**) and DM (ρ = 0.72, **Fig. 2b**). Subsequently, we integrated clinical and proteomics data through DISCO analysis, referred to as ClinicalProt. The correlation of ClinicalProt DISCO with age became even stronger (ρ = 0.48, **Fig. 2d**) while its correlation with DM decreased (ρ = 0.63, **Fig. 2e**). DISCO's performance in capturing mortality risks (C-index = 0.741) surpassed that of DM (C-index = 0.677). By categorizing individuals into low-risk and high-risk groups based on DISCO quartiles, the survival differences between the two groups were more pronounced than those observed with DM (**Fig. 2c, f, Supp fig. 2**).

In the RuLAS study (Wave 4), by using 2-year mortality as the outcome measure, we found that DISCO values derived from all clinical biomarkers effectively identified individuals at high risk (**Fig. 2g**), outperforming age as a predictor (C-index: DISCO = 0.684; age = 0.642). Furthermore, the dysregulation of metabolomics homeostasis, as measured by DISCO analysis on NMR metabolomics data, was associated with an increased risk of 2-year mortality after

adjusting for age and sex (aHR: 1.17, 95% CI: 1.06-1.29, C-index: 0.578). We identified 27 microbial species correlated with host metabolism homeostasis (Spearman $p < 0.05$) and constructed a DISCO score for the microbiota. The dysregulation of gut microbial homeostasis demonstrated a stronger association with 2-year mortality risks (aHR: 1.38, 95% CI: 1.10-1.74, C-index: 0.610). When integrating microbiota and metabolic homeostasis with organ systems, the aggregate DISCO score significantly predicted mortality (aHR: 2.49, 95% CI: 1.92-3.24, C-index: 0.713), surpassing the predictive power of single data type DISCO.

Validation of DISCO's robustness across reference groups, biomarker subsets, and distance metrics confirmed stable performance (**Extended Fig. 2**). The DISCO distribution held relatively stable across the reference group sample size and biomarker selection (**Extended Fig. 2a-b**). When selecting the young adults as the reference, the prediction power for mortality reached the highest C-index, and gradually decreased as the reference sample age increased (**Extended Fig. 2c**). This was due to the increase of homeostatic dysregulation with age. The DISCO based on various distance metrics yielded similar prediction performance for mortality (**Extended Fig. 2d**). When the biomarker numbers increased, the superiority of DISCO over DM approach became distinct (>= 20 biomarkers) and stabilized with enough biomarkers (>=100), dependent on the biomarker data types (**Extended Fig. 2e**). Leave-one-out analysis (**Extended Fig. 2f**) indicated the robustness of DISCO on non-specificity for a single biomarker.

DISCO predicted the incidences of clinical outcomes

Next, we investigated the associations between DISCO related to clinical biomarkers (Clinical), proteins (Prot), and their combined sets (ClinicalProt) with the incidences of age-related chronic diseases in the UKB (**Fig. 3a**). The chronic diseases included obesity, hypertension, diabetes, cancer, myocardial infarction (MI), stroke, chronic obstructive pulmonary disease (COPD), and dementia. DISCO was found to be predictive of these chronic conditions, with DISCO ClinicalProt exhibiting the highest prediction accuracy for risk capture, followed by DISCO based on proteins and clinical biomarkers. Additionally, analysis of the same datasets revealed that the higher C-index of DISCO indicated its superiority over DM in capturing disease risks, such as diabetes (DISCO C-index: 0.721; DM C-index: 0.669). The disease Kaplan-Meier curves (**Extended Fig. 3a, Supp Fig. 3-4)** demonstrated that the highest and lowest quartiles of DISCO exhibited more pronounced differences compared to those observed with DM. These findings highlight the potential of DISCO to integrate clinical and omics data to reflect dysregulation in homeostasis. Furthermore, it appears that homeostatic dysregulation occurs prior to disease diagnosis and is capable of predicting subsequent incidence of diseases.

In the CHARLS study, we explored the relationship between DISCO and frailty. The frail group exhibited the highest DISCO levels, followed by the prefrail and robust groups (**Extended Fig. 3b**). During the longitudinal follow-up from Wave 1 to Wave 3 (**Fig. 3b**), higher DISCO levels were associated with an increased risk of developing frailty, with an odds ratio (OR) of 1.11 (95% CI: 1.07-1.16). This association was validated in newly enrolled participants during CHARLS Waves 3-4, with an OR of 1.20 (95% CI: 1.05-1.37). When jointly assessing DISCO

levels and frailty status (**Fig. 3c**), even robust individuals with high DISCO were at higher risk of mortality than frail individuals with low DISCO. Moreover, we evaluated the predictive accuracy of DISCO and the frailty index in forecasting binary mortality outcomes. The area under the curve (AUC) for DISCO was 0.702 and 0.729 at the two follow-ups, while the AUC for the frailty index was 0.667 and 0.693 (**Extended Fig. 3c**). Compared with DM approach, the ROC curve analysis indicated that DISCO provided more accurate mortality predictions than DM and added substantial value above and beyond age. For example, in the RuLAS cohort, the area under the curve (AUC) values for age, DISCO, and DM were 0.648, 0.738, and 0.710, respectively. Note that the differences in the C-statistic of age across cohorts are likely due to differences in the age range of these cohorts.

DISCO characterized homeostatic dysregulation of organs and tissues

Using the UKB Olink proteomics panel, we derived DISCO scores for 18 organs and tissues (e.g., brain, liver, kidney, intestine) based on enriched organ-specific proteins (**Fig. 4a, Extended Fig. 4a-c**). Clinical biomarkers showed weak-to-moderate correlations with their corresponding organ DISCO scores (e.g., aspartate aminotransferase [AST] with liver DISCO). After adjusting for age and sex, cystatin C displayed stronger associations with DISCO scores across multiple organs (**Fig. 4b**). Organ-specific DISCO values exhibited moderate age correlations, with intestine, brain, and kidney showing the highest correlations (**Extended Fig. 4d**). Number of organ/tissue DISCO scores in the top 10% increased exponentially with age fitting through survival analysis (**Fig. 4c, Extended Fig. 4e**), mirroring mortality risk trends. Organ-specific risk profiles revealed both age-dependent progression and organ/tissue-specific variability (**Fig. 4d**). Sensitivity analysis (**Extended 4f-g**) confirmed organ DISCO's associations with mortality, and its high-risk exponential increase of high-risk DISCO organ and organ/tissue-specific patterns with age.

In terms of clinical outcomes, the DISCO values for the intestine, brain, kidney, choroid plexus, salivary gland, lung, and lymphoid tissue all had C-index values exceeding 0.65 in predicting all-cause mortality. Elevated levels of organ-specific DISCO were linked to chronic/organ-related diseases (**Fig. 4e**), as well as to mortality from these diseases. There was some specificity of the organ-specific DISCO score and the relevant diseases. For example, only brain DISCO predicted Alzheimer's disease mortality. However, globally such specificity was surprisingly weak. Heart disease and respiratory disease mortality were predicted more strongly by every single organ-specific DISCO score than any other cause of mortality, and lung DISCO only barely outperformed brain, intestine, and salivary DISCOs to predict respiratory disease mortality. Heart muscle DISCO was only the 5th-best predictor of heart disease mortality. These findings suggest systemic homeostatic dysregulation influences disease risks more prominently than organ-specific dysfunction alone.

In RuLAS (Wave 4), we examined homeostatic dysregulation of organ systems using organ-specific biomarkers such as heart (e.g., brain natriuretic peptide), kidney (e.g., cystatin C), vascular health (e.g., homocysteine), oxygen transport (e.g., hemoglobin), and immunity (e.g.,

IL-6), among others. Organ-specific biomarkers, related to vascular health, oxygen transport, and kidney function exhibited the highest predictive capacity for 2-year mortality (**Fig. 4f**). Each organ DISCO contributed a marginal contribution to the mortality, and the aggregate DISCO had the highest predictive capacity for mortality, indicating multi-organ homeostatic dysregulations.

Interconnection among homeostatic dysregulation of organs and tissues

Correlation network analysis of clinical and proteomics biomarkers with ClinicalProt DISCO (adjusted for age and sex) revealed interconnected hubs, including GDF15, TNFRSF10B, and EDA2R (**Fig. 5a-b**). Furthermore, DISCO effectively captured key aging-related information in proteomic data, showing a robust correlation ($\rho = 0.74$; **Supp Fig. 5**) that was significantly higher than those derived from DM. At the organ level, DISCO values demonstrated significant interconnections (**Fig. 5c**), with age- and sex-adjusted partial correlation coefficients spanning 0.3 to 0.7, highlighting coordinated dysregulation patterns across organs. Among these organs and tissues, the brain, intestine and esophagus had the highest degree on the K-neighbor network (**Fig. 5d**), suggesting these organs had greater influence on global homeostatic regulation. The degree centrality of organs' homeostatic dysregulations showed correlations with their Pearson's correlation with age, C-index for mortality prediction, and mean C-index for chronic disease prediction (**Fig. 5e, Extended Fig. 5a**). The selection of weighted power of organs' partial correlation did not affect the significance of these results (**Extended Fig. 5b**). Sensitivity analysis (**Extended Fig. 5c-e**) showed the selection of the same number of proteomics biomarkers in DISCO analysis with replicates would also yield consistent results of organs' homeostatic dysregulation interconnection. In summary, the DISCO values for various organs and tissues were interconnected and predictive of mortality and disease risks beyond their specific organ correspondence.

Discussion

In this study, we developed and validated a new approach to measuring homeostatic dysregulation, DISCO. DISCO generally performed better at predicting health outcomes than the best-known measure of homeostatic dysregulation, D_M. It was moderately correlated with age but strongly predicted mortality, the frailty index, and every chronic condition tested (aHR often greater than 1.5 per unit). DISCO is an unsupervised method and is not specifically trained to predict anything; it nonetheless predicted mortality comparably to the best epigenetic clocks, which were specifically trained to do so[20,21]. DISCO also performed well on other types of data, including proteomics, metabolomics, and microbiomes, substantially outperforming D_M as the number of biomarkers increased. Furthermore, when we generated organ- and tissue-specific DISCO scores based on circulating proteins known to be associated with the target organ/tissue, we observed only very limited specificity of the prediction of a given disease or cause of death based on the organ/tissue most obviously associated. For example, lung DISCO predicted respiratory disease mortality only slightly better than brain, intestine, or salivary DISCO, and

heart muscle DISCO was the fifth-best at predicting heart disease mortality, after brain, intestine, liver, and adipose tissue DISCO. Network analyses showed that well-known aging associated proteins were both central in networks and strongly associated with DISCO; furthermore, organ/tissue DISCOs most central in networks were those that most strongly associated with age and health outcomes.

As a measure of dysregulation, the superiority of DISCO to D_M was unequivocal. Prediction of mortality was similar between the two when applied to standard clinical biomarkers, but DISCO was consistently more predictive when applied to proteomics, microbiomes, or combinations of data types. Furthermore, it provided substantially stronger prediction of incidence of specific diseases and chronic conditions than D_M and showed excellent performance across datasets. In comparison to epigenetic clocks in NHANES, not only were C-indices for mortality comparable with the best clocks; the slightly weaker association of DISCO with age meant that DISCO had substantially better prognostic value at younger ages (**Fig. 1h**). Even in comparison with the frailty index, which is known to be a strong predictor of health outcomes[22,23], DISCO proved to be a stronger discriminator of mortality risk, with high-risk DISCO individuals having greater risk of mortality than low-risk individuals regardless of frailty status (frail, prefrail and robust; **Extended Fig. 3c**). In other words, it is better to be frail with low DISCO than robust with high DISCO. While DISCO has not yet been validated in the wide variety of contexts these other measures have, our findings here suggest that DISCO is likely to be the best broad predictor yet identified of health outcomes during aging.

Perhaps our most striking finding was that organ/tissue DISCO scores were broadly and strongly predictive of specific health outcomes, irrespective of the obvious expected associations between the organ/tissue and the health outcome. Rather, certain outcomes were broadly more sensitive to DISCO than others (e.g., respiratory and heart disease mortality was strongly predicted, with HRs often well above 2, whereas Alzheimer's disease was weakly associated, if at all). Likewise, some organ/tissue DISCO scores were broadly more predictive than others: brain, intestine, lung, lymphoid tissue, and salivary gland showed stronger associations with most outcomes, whereas retina, tongue, and stomach associations were weaker. The organ/tissue DISCOs with the strongest effects were also those that were most central in networks of these scores – i.e., those most broadly correlated with others. This finding was robust across sensitivity analyses changing the inclusion criteria for proteins within a DISCO score and adjusting the number of proteins included. It contrasts with findings in D_M showing that system-specific D_Ms are only weakly correlated after age adjustment[12].

Our approach to measuring organ-specific dysregulation based on circulating plasma proteins differs from other approaches to measuring organ/system aging via blood. Like us, Oh et al.[24], Goeminne et al.[25] and Kivimäki et al.[26] use plasma proteomics with proteins assigned to organs via the Human Proteome Atlas. They then generate organ aging scores as the difference between observed and predicted chronological age, with the prediction via LASSO or elastic net

regression trained on age and/or mortality. In predicting health outcomes, their results are more organ-specific than ours, though not perfectly so. For example, in predicting congestive heart failure, Oh et al.'s heart organ age gap was a substantially stronger predictor than any other organ age gap, and only 5 of the 13 organs were significant predictors. Furthermore, organ/tissue DISCO scores appear to be substantially more strongly associated with health outcomes than Oh's organ age gaps: the strongest organ age gap predictor of mortality was HR=1.54; 12 of 18 organ/tissue DISCOs had stronger HRs than this. Because Goeminne et al. trained on mortality rather than age, their HRs are stronger than for Oh et al., with their strongest HRs comparable to ours. Their results are somewhat less organ specific than Oh et al. and Kivimäki et al., but more so than ours. Seghal et al. use DNA methylation to construct system-specific aging clocks[27]. They observe substantially more organ-disease specificity than we do; relative performance of the approaches is hard to evaluate due to differing effect size metrics. Overall, these other studies suggest that supervised machine learning approaches will identify organ/system scores that are more targeted in their effects than the dysregulation approaches we used. We also note that our unsupervised approach provides predictions at least as strong as approaches specifically trained to predict aging/mortality.

Levels of individual biomarkers often do have specific associations with certain chronic conditions – for example, cystatin C is thought to be relatively specific to kidney disease[28]. Nonetheless, here cystatin C was also the single biomarker with the largest impact on DISCO when excluded. Dysregulation scores are qualitatively different than linear or additive biomarker scores because they can be agnostic to precise levels of the biomarkers, asking instead about the joint configuration of many biomarkers and how that configuration deviates from homeostatic norms. Our findings strongly suggest that such dysregulation scores have relatively little specificity to the tissue of origin of relevant biomarkers, and that they rather detect global, organism-level coordination. In the context of aging research, this implies that the coordination across tissues and organs is at least as important as organ-specific processes. This is consistent with studies on dynamical inter-organ signaling during aging[29,30], as well as with studies showing broad collapse across physiological systems in the months preceding death in hemodialysis patients[31,32]. In the context of research on specific chronic conditions, this suggests we need to think beyond what is observed in the pathways and tissues within the organ: key aspects of function are coordinated with the rest of the body, and understanding disease in an organ is likely to require integration with external condition and signaling.

At a molecular, level, DISCO scores were particularly strongly associated with several proteins known to play key roles in aging, including GDF15, AGRN, and some of the TNF receptor superfamily[33-35]. GDF15 in particular has repeatedly emerged as a strong predictor of aging, psychological stress, and energetic stress[36-39]. It is triggered at a cellular level by the integrated stress response and lack of energy availability[40-42]. This suggests that DISCO may be at least partially quantifying the breakdowns in energetic efficiency that occur for a myriad of reasons during aging, and that manifest as disruptions and adjustments in regulatory networks. The tight

linkage between DISCO and the aging process as a whole is supported by the nearly perfect alignment between Gompertzian increase in mortality and the accrual high-risk systems in DISCO. Several lines of evidence thus suggest that DISCO is detecting fundamental manifestations of the aging process and that aging is tightly aligned with loss of control in the complex dynamic networks that maintain organismal homeostasis.

Future work should test the feasibility of DISCO as a general measure of health and aging in diverse contexts, such as its responsiveness to upstream determinants of aging and health (e.g., psychological stress, lifestyle, and environmental) as well as its more specific prognostic value for health outcomes under a range of circumstances. The utility of individual organ/tissue DISCO scores remains to be seen, given their strong intercorrelations; potentially, a more robust global DISCO might be generated by appropriately integrating the organ/tissue scores than by calculating directly. Future methodological work should also focus on more precisely defining the sensitivity of DISCO to the reference population used and, ideally, work toward development of a universal reference population that would allow standardized calculation in any dataset[43]; this may be challenging, given that DISCO relies on an entire correlation matrix. Nonetheless, calculation of DISCO is straightforward within a dataset of sufficient size (N > 200), and we expect it to be a readily accessible metric for use in population research on aging.

Our findings are consistent with, but do not fully prove, the broad hypothesis that aging emerges holistically from numerous mechanisms and causes operating at different biological scales and creating dynamic feedback loops, rather than from a limited set of discrete mechanisms. Under this hypothesis, organisms are complex dynamic systems that have evolved multiple repair and error-correction mechanisms to resist entropy (the second law of thermodynamics) as long as possible[1,3] . However, these mechanisms are imperfect, and entropy can accrue in many ways, from DNA damage to changes in tissue structure to imbalances among signaling molecules. The patterns of DISCO observed here are what would be expected under this entropy hypothesis, and indeed DISCO could be considered an individual-level approximation of entropy; however, we have not here distinguished whether this entropy can be traced upstream to a more limited set of causes. Future work could explore the extent to which DISCO is or is not largely determined by canonical aging mechanisms such as the hallmarks[44].

Methods and materials

Study populations

Our study used the data of NHANES 1999-2018, UKB, CHARLS, CLHLS, and RuLAS. The US NHANES is a nationally representative cross-sectional survey of civilian living in the US, approved by the National Center for Health Statistics (NCHS) Ethics Review Board[45]. The UK Biobank is large-scale prospective cohort that collected data from over 500,000 participants across 22 centers in England, Scotland, and Wales. UKB received ethics approval from the North West Multicenter Research Ethics Commitee[46]. The CHARLS is an ongoing prospective population-based longitudinal cohort study of middle-aged and older Chinese adults. CHARLS

was approved by the Ethics Review Board of Peking University, which was conducted in accordance with the Declaration of Helsinki and other relevant guidelines and regulations[47]. The CLHLS is a nationwide longitudinal study of old-aged Chinese population. The project was approved by the Biomedical Ethics Committee of Peking University, China (IRB00001052-13074)[48]. The Rugao Longevity and Ageing Study (RuLAS) is a population-based prospective study, which consisted of a longevity cohort and an aging cohort in Rugao, China[16]. The RuLAS was approved by the Human Ethics Committee of Fudan University School of Life Sciences. All participants provided written informed consent. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines for cohort studies.

Collection of clinical biomarkers

Here, we use as many clinical biomarkers as possible for each cohort, based on what is accessible. In NHANES, 24 biomarkers of blood biochemical assays (glucose, albumin, etc.) and complete blood count tests (NEU%, LYM%, etc.) were collected. In UKB, 58 biomarkers of hematological assays and biochemistry tests measured in the blood sample were collected in this study. In RuLAS, fasting blood samples were collected for laboratory tests (physical, hematological, and biochemical biomarkers). In RuLAS, 57 biomarkers of Wave 2 (2016) and 63 biomarkers of Wave 4 (2019) were enrolled for analysis, respectively. For CHARLS, both the Wave 1 (2011) and Wave 3 (2015) collected blood samples for complete blood count (CBC) tests and blood-based bioassays, in which 16 biomarkers were measured for the participants. For CLHLS (2014 wave), 30 biomarkers of CBC, blood biochemical examination and urine routine test were involved in DISCO analysis. The details of biomarkers were summarized in **Supplementary Tables 1-6**.

Collection of proteomics biomarkers

In UKB (2006-2010), the raw plasma proteomics data (Olink platform) contained 2923 proteins from 53,014 participants. In quality control, we removed proteins with more than 10% missing data and participants with more than 50% missing data, resulting in 1461 protein biomarkers on 47,859 participants for analysis. To reduce the computational cost and remove unconcerned variables, we used LASSO-cox regression model (five-fold cross validation) to extract proteins related to mortality risk. We selected the penalty as the model reached the highest C-index for survival analysis. As a result, 367 proteins were obtained through five-fold cross-validation. For the union analysis, 47,173 participants with 58 clinical biomarkers and 367 proteins were enrolled. After feature selection, 396 features (39 biomarkers and 357 proteins) were retained (**Supplementary Tables 7-8**).

Organ-specific proteomics analysis

The organ and tissue enriched and enhanced proteins were obtained from the Human Protein Altas (https://www.proteinatlas.org). Consequently, proteins from the proteomics analysis were assigned as appropriate to one of 36 tissues. We selected the organs/tissues with more than 40

such proteins for DISCO analysis, and sex-related organs were removed (placenta, ovary, testis). Following the feature selection in previous studies[24,25], the organ/tissue specific and enhanced proteins were pruned using Lasso-cox regression models with the penalty as the model reached highest C-index (**Supplementary Table 9**). To define organ-specific high-risk DISCO, we employed quantile-based thresholds to estimate parameters of the Gompertz-Makeham (GM) model, which captures age-dependent increases in dysregulation risk. GM parameters stabilized at thresholds ≥10% (Extended Fig. 4e), establishing this as the critical cutoff. We then computed an adjusted "number at risk" by dividing the observed age- and individual-averaged number of high-risk DISCO organs by the expected value under uniform distribution assumptions. The expected value was defined as the total number of organs multiplied by the 10% threshold. For pairwise correlation analysis of inter-organ DISCO levels, we first removed overlapping proteins to eliminate redundancy and recalculated DISCO values along with their Pearson correlation coefficients. This preprocessing step ensured that the analysis focused on unique contributors to homeostatic dysregulation across organs. In addition, we performed a comprehensive sensitivity analysis through the following protocol: (1) Protein Sampling. First, organs/tissues were ranked in ascending order based on their protein counts. Following this ranking, 40 proteins were randomly selected from each organ/tissue, a process repeated across 200 independent iterations across 200 independent replicates; (2) Non-overlap constraint. Proteins were sampled under a strict exclusion rule to prevent inter-sample overlap, ensuring independence of observations. DISCO values were computed for each replicate and each organ/tissue. Finally, we performed global correlation analysis on the mean DISCO profiles to identify significant correlations between organs/tissues.

Collection of metabolomics and microbiota biomarkers

In RuLAS Wave 4 (2019), we collected blood serum metabolomics (profiling through nuclear magnetic resonance) and gut microbiome data[49]. For metabolomics data, 313 metabolites and related variables (amino acid, low-density and high-density lipoproteins, derived ratios, etc) from 2120 participants were collected. Due to strong correlation and collinearity of the metabolomics data, we used principal components analysis to reduce the redundancy and extract the primary information. The first 50 principal components explaining over 95% of the variance of the original variables, were used for DISCO analysis (**Supplementary Table 10**). For gut microbiome data, we selected the microbiota with missing abundancy less than 95% across all participants. Next, 251 microbiota of 1525 participants were used for analysis. We used the metabolomics data to measure the metabolic homeostatic dysregulation, and 27 microbiotas correlated with metabolic homeostasis were selected (Spearman p value < 0.05, **Supplementary Table 11**). DM analyses followed the same analytical workflow as DISCO, including biomarker selection, dimensionality reduction, and statistical significance. A total of 30 microbiotas correlated with metabolic homeostasis (defined by DM) were selected, which

were highly consistent with those of DISCO with 26 microbiotas overlapped. DM analyses adhered to the same analytical workflow as DISCO, encompassing dimensionality reduction, biomarker selection, and statistical validation. A total of 30 microbiota associated with metabolic homeostasis (as defined by DM) were identified, with 26 overlapping taxa (86.7% consistency) shared between DM and DISCO.

The calculation of DISCO

The Euclidean Distance of Covariance (DISCO) was defined as the squared sum of differences between the elements of the perturbed covariance matrix and the reference covariance matrix. Assuming that there were m biomarkers, DISCO was derived using Pearson correlation coefficients (PCCs) from two groups: a reference group with n samples (denoted as ref_n) and a perturbed group with n+1 samples (denoted as ref_{n+1}). Then, we added a weighted score to the component elements of the DISCO, aiming to enhance the weight of aging biomarkers. The DISCO metric is calculated as:

$$DISCO = \mathrm{n}^2 * \sum_{i=1}^{m} \sum_{j=1}^{i-1} Weight_{i,j} (PCC_{i,j}^{(ref_{n+1})} - PCC_{i,j}^{ref_n})^2$$

and $Weight_{i,j} = \frac{PCC_{biomarker_i,age} * PCC_{biomarker_j,age}}{\sum\sum PCC_{biomarker_i,age} * PCC_{biomarker_j,age}}$

The $\sum\sum PCC_{biomarker_i,age} * PCC_{biomarker_j,age}$ was the absolute value of quadratic sum of PCC between biomarker i, j, and age, working as a scale factor. The raw distribution of DISCO was right-skewed, and we used log transformed DISCO in downstream analysis.

Assessment of clinical outcomes

In NHANES, death information was based on linked data from records taken from the National Death Index (NDI) through December 31, 2019, provided through the Centers for Disease Control and Prevention. In UKB, death information was obtained through death certificates held within the National Health Service (NHS) Information Centre (England and Wales) and the NHS Central Register (Scotland) to November 30, 2022. We used the International Statistical Classification of Diseases, 10th, to define causes of death. The cause-specific mortality included mortality of malignant neoplasm, heart disease, cerebrovascular disease, respiratory disease, Alzheimer disease, diabetes, and others. In addition, diagnosed dates of incident chronic disease in UKB were also collected, including obesity, hypertension, diabetes, cancer, myocardial infraction, heart failure, stroke, chronic obstructive pulmonary disease (COPD), and dementia. In RuLAS, dates of death for the participants were obtained from the Bureau of Civil Affairs of Rugao. For RuLAS Wave 2 (2016) and Wave 4 (2019), 4-year survival and 2-year survival data were calculated, respectively. In CHARLS, records on death were obtained at the biennial follow-up. Because the exact date of death was not available, the point of the last follow-up visits was defined as the time of death. In CHARLS, we used frailty

index (FI) to evaluate the frailty status using chronic diseases and functional limitations were collected for frailty assessment (**Supplementary Table 12**). Participants were divided into robust (FI<0.1), prefrail (0.1<=FI<0.25) and frail groups (FI>=0.25). In CLHLS, death information was collected from the official death certificates when available. Otherwise, the information was obtained from next-of-kin and local residential committees.

Benchmark and robustness analysis

DISCO's predictive performance for mortality and disease outcomes was benchmarked against DM. We further analyzed DNA methylation aging clock data (**Supplementary Table 13**) from the NHANES 1999–2002 cohort, which initially included 4,449 participants. After excluding 1,917 individuals with incomplete methylation data, 2,532 participants were retained for analysis. Over a median follow-up of 17.2 years (IQR: 9.3–18.8), 1,361 all-cause mortality events were recorded. All-cause mortality served as the outcome for comparing DISCO and DM against established aging clocks. To assess DISCO's robustness, we evaluated its distribution stability and feature correlation under varying reference group sizes and biomarker selection strategies, conducting 50 randomized replicates for each parameter configuration. We further examined its performance across different age-stratified reference group definitions and alternative matrix distance metrics. Additionally, DISCO's predictive capacity for mortality was systematically tested with incremental biomarker inclusion, with mean performance metrics calculated over 50 randomized replicates at each step. Finally, leave-one-biomarker-out analysis validated its robustness to individual feature selection, with performance distributions estimated via 50 bootstrap iterations.

Statistical analysis

Participants were classified into quartiles through DISCO/DM, with the first and fourth quartiles representing individuals at lowest and highest risks, respectively. Kaplan-Meier survival curves were then plotted to compare the predicted survival probabilities. We utilized Cox proportional hazard models to estimate the hazard ratios (HRs) and 95% confidence intervals (CIs) of levels associated with mortality/diseases using two models. Model 1: crude model; Model 2: adjusted for age and sex. The association analysis was shown in **Supplementary Tables 14-16**. From the Cox models, each individual's linear predictor was then used to calculate the concordance index (C-index), a measure of the model's discriminative ability. Harrell's C-index is known to be biased in the presence of right censoring[50], which is present in our datasets. Although Uno's C-index[51] corrects for this bias, it relies on the assumption of non-informative censoring, which may not hold in practice. We tested both indices for a subset of analyses (mortality prediction in NHANES, **Fig. 1g** and **Supp Fig. 1d**) and found nearly identical results. Accordingly, we used Harrell's C-index throughout to assess the predictive discrimination in survival analysis. For binary outcomes (x-year moratlity, x varied by outcomes), the Area Under the Curve (AUC) was took as a robust metric to evaluate the prediction ability. Cox proportional hazard models were conducted to assess the

associations between different biological aging clocks and the onset of chronic diseases, including chronic obstructive pulmonary disease (COPD), dementia, and others. These Cox models were adjusted for age and sex. All statistical analyses were performed using R version 4.3.3.

Acknowledgments

The data used in this research were obtained from the NHANES, UKB, CHARLS, RuLAS and CLHLS. We would like to thank the workers, researchers, and participants involved in these cohorts.

Funding

This work was supported by grants from the National Natural Science Foundation of China-Youth Science Fund (32300533, 82301768, 32100510), the Shanghai Sailing Program (23YF1430500).

Conflict of interest

None declared.

Data sharing statement

The data of RulAS are available through reasonable request from the corresponding author. The data from CLHLS are available at https://opendata.pku.edu.cn/dataset.xhtml. The data from CHARLS are available at https://charls.charlsdata.com/pages/data/111/zh-cn.html. The data from the NHANES are available at www.cdc.gov/nchs/nhis/index.htm, and the data from the UK Biobank are available upon application at www.ukbiobank.ac.uk/register-apply. This research was conducted using UK Biobank Resource under Application Number 103791.

Author contributions

Concept and design: Meng Hao, Hui Zhang.

Acquisition, analysis, or interpretation of data: Meng Hao, Alan A Cohen, Yi Li, Zixin Hu, Shuai Jiang, Kamaryn T. Tanner.

Drafting of the manuscript: Meng Hao, Hui Zhang, Alan A Cohen.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Meng Hao, Hui Zhang, Yi Li, Yaqi Huang, Kamaryn T. Tanner.

Administrative, technical, or material support: Xiangnan Li, Yaqi Huang, Shuishan Zhang, Yiqin Huang, Jie Chen, Zhijun Bao, Jiucun Wang, Li Jin.

Supervision: Xiaofeng Wang, Alan A Cohen, Yiqin Huang, Jiucun Wang.

Reference

1 Tatullo, M. Entropy Meets Physiology: Should We Translate Aging as Disorder? *Stem cells (Dayton, Ohio)* **42**, 91-97, doi:10.1093/stmcls/sxad084 (2024).

2 Wensink, M. J. & Cohen, A. A. The Danaid Theory of Aging. *Frontiers in cell and developmental biology* **9**, 671208, doi:10.3389/fcell.2021.671208 (2021).

3 Bortz, W. M., 2nd. Aging as entropy. *Experimental gerontology* **21**, 321-328, doi:10.1016/0531-5565(86)90039-2 (1986).

4 Schrodinger, E. What is life?: the physical aspect of the living cell. (1946).

5 Ferrucci, L. & Kuchel, G. A. Heterogeneity of Aging: Individual Risk Factors, Mechanisms, Patient Priorities, and Outcomes. *J Am Geriatr Soc* **69**, 610-612, doi:10.1111/jgs.17011 (2021).

6 Lowsky, D. J., Olshansky, S. J., Bhattacharya, J. & Goldman, D. P. Heterogeneity in healthy aging. *The journals of gerontology. Series A, Biological sciences and medical sciences* **69**, 640-649, doi:10.1093/gerona/glt162 (2014).

7 Kedlian, V. R., Donertas, H. M. & Thornton, J. M. The widespread increase in inter-individual variability of gene expression in the human brain with age. *Aging* **11**, 2253-2280, doi:10.18632/aging.101912 (2019).

8 Franceschi, C. *et al.* Immunobiography and the Heterogeneity of Immune Responses in the Elderly: A Focus on Inflammaging and Trained Immunity. *Frontiers in immunology* **8**, 982, doi:10.3389/fimmu.2017.00982 (2017).

9 Kraft, T. S. *et al.* Multi-system physiological dysregulation and ageing in a subsistence population. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **375**, 20190610, doi:10.1098/rstb.2019.0610 (2020).

10 Senior, A. M. *et al.* Multidimensional associations between nutrient intake and healthy ageing in humans. *BMC biology* **20**, 196, doi:10.1186/s12915-022-01395-z (2022).

11 Cohen, A. A. *et al.* A novel statistical approach shows evidence for multi-system physiological dysregulation during aging. *Mech Ageing Dev* **134**, 110-117, doi:10.1016/j.mad.2013.01.004 (2013).

12 Li, Q. *et al.* Homeostatic dysregulation proceeds in parallel in multiple physiological systems. *Aging Cell* **14**, 1103-1112, doi:10.1111/acel.12402 (2015).

13 Lu, Y. *et al.* Physiological Dysregulation, Frailty, and Impacts on Adverse Health and Functional Outcomes. *Frontiers in medicine* **8**, 751022, doi:10.3389/fmed.2021.751022 (2021).

14 Egorov, A. I. *et al.* Greater tree cover near residence is associated with reduced allostatic load in residents of central North Carolina. *Environmental research* **186**, 109435, doi:10.1016/j.envres.2020.109435 (2020).

15 Jin, S. *et al.* Association of lifestyle with mortality and the mediating role of aging among older adults in China. *Archives of gerontology and geriatrics* **98**, 104559, doi:10.1016/j.archger.2021.104559 (2022).

16 Hao, M. *et al.* Phenotype correlations reveal the relationships of physiological systems underlying human ageing. *Aging Cell* **20**, e13519, doi:10.1111/acel.13519 (2021).

17 Li, Y. *et al.* The composite phenotype analysis identifies potential concerted responses of physiological systems to high altitude exposure. *Natl Sci Rev* **10**, nwad053, doi:10.1093/nsr/nwad053 (2023).

18 Hao, M. *et al.* Metrics of Physiological Network Topology Are Novel Biomarkers to Capture Functional Disability and Health. *J Gerontol A Biol Sci Med Sci* **80**, doi:10.1093/gerona/glae268 (2024).

19 Hao, M. *et al.* Using Physiological System Networks to Elaborate Resilience Across Frailty States. *J Gerontol A Biol Sci Med Sci* **79**, doi:10.1093/gerona/glad243 (2024).
20 Lu, A. T. *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303-327, doi:10.18632/aging.101684 (2019).
21 Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* **10**, 573-591, doi:10.18632/aging.101414 (2018).
22 Rockwood, K. *et al.* A Frailty Index Based On Deficit Accumulation Quantifies Mortality Risk in Humans and in Mice. *Scientific reports* **7**, 43068, doi:10.1038/srep43068 (2017).
23 Rockwood, K. & Mitnitski, A. Frailty in relation to the accumulation of deficits. *The journals of gerontology. Series A, Biological sciences and medical sciences* **62**, 722-727, doi:10.1093/gerona/62.7.722 (2007).
24 Oh, H. S. *et al.* Organ aging signatures in the plasma proteome track health and disease. *Nature* **624**, 164-172, doi:10.1038/s41586-023-06802-1 (2023).
25 Goeminne, L. J. E. *et al.* Plasma protein-based organ-specific aging and mortality models unveil diseases as accelerated aging of organismal systems. *Cell Metab* **37**, 205-222 e206, doi:10.1016/j.cmet.2024.10.005 (2025).
26 Kivimaki, M. *et al.* Proteomic organ-specific ageing signatures and 20-year risk of age-related diseases: the Whitehall II observational cohort study. *Lancet Digit Health* **7**, e195-e204, doi:10.1016/j.landig.2025.01.006 (2025).
27 Sehgal, R. *et al.* Systems Age: A single blood methylation test to quantify aging heterogeneity across 11 physiological systems. *bioRxiv*, 2023.2007.2013.548904, doi:10.1101/2023.07.13.548904 (2024).
28 Benoit, S. W., Ciccia, E. A. & Devarajan, P. Cystatin C as a biomarker of chronic kidney disease: latest developments. *Expert review of molecular diagnostics* **20**, 1019-1026, doi:10.1080/14737159.2020.1768849 (2020).
29 Romero-Ortuño, R. *et al.* Network Physiology in Aging and Frailty: The Grand Challenge of Physiological Reserve in Older Adults. *Frontiers in network physiology* **1**, 712430, doi:10.3389/fnetp.2021.712430 (2021).
30 Garcia-Retortillo, S. *et al.* Age-related breakdown in networks of inter-muscular coordination. *GeroScience* **47**, 1615-1639, doi:10.1007/s11357-024-01331-9 (2025).
31 Nakazato, Y. *et al.* Estimation of homeostatic dysregulation and frailty using biomarker variability: a principal component analysis of hemodialysis patients. *Scientific reports* **10**, 10314, doi:10.1038/s41598-020-66861-6 (2020).
32 Cohen, A. A. *et al.* Synchrony of biomarker variability indicates a critical transition: Application to mortality prediction in hemodialysis. *iScience* **25**, 104385, doi:10.1016/j.isci.2022.104385 (2022).
33 Pratt, J. *et al.* Genes encoding agrin (AGRN) and neurotrypsin (PRSS12) are associated with muscle mass, strength and plasma C-terminal agrin fragment concentration. *Geroscience* **45**, 1289-1302, doi:10.1007/s11357-022-00721-1 (2023).
34 Conte, M. *et al.* GDF15, an emerging key player in human aging. *Ageing Res Rev* **75**, 101569, doi:10.1016/j.arr.2022.101569 (2022).
35 Kremers, B. M. M. *et al.* Discovery of four plasmatic biomarkers potentially predicting cardiovascular outcome in peripheral artery disease. *Sci Rep* **12**, 18388, doi:10.1038/s41598-022-23260-3 (2022).

36 Huang, Q. *et al.* The Energetic Stress Marker GDF15 is Induced by Acute Psychosocial Stress. *bioRxiv : the preprint server for biology*, doi:10.1101/2024.04.19.590241 (2024).

37 Shen, C. *et al.* Plasma proteomic signatures of social isolation and loneliness associated with morbidity and mortality. *Nature human behaviour* **9**, 569-583, doi:10.1038/s41562-024-02078-1 (2025).

38 Evans, D. S. *et al.* Proteomic Analysis of the Senescence-Associated Secretory Phenotype: GDF-15, IGFBP-2, and Cystatin-C Are Associated With Multiple Aging Traits. *The journals of gerontology. Series A, Biological sciences and medical sciences* **79**, doi:10.1093/gerona/glad265 (2024).

39 Tanaka, T. *et al.* Plasma proteomic biomarker signature of age predicts health and life span. *eLife* **9**, doi:10.7554/eLife.61073 (2020).

40 Shaulson, E. D., Cohen, A. A. & Picard, M. The brain-body energy conservation model of aging. *Nature aging* **4**, 1354-1371, doi:10.1038/s43587-024-00716-x (2024).

41 Asundi, J. *et al.* GDF15 is a dynamic biomarker of the integrated stress response in the central nervous system. *CNS neuroscience & therapeutics* **30**, e14600, doi:10.1111/cns.14600 (2024).

42 Smith, J. L. *et al.* Mapping the GDF15 Arm of the Integrated Stress Response in Human Cells and Tissues. *bioRxiv : the preprint server for biology*, doi:10.1101/2025.01.31.635929 (2025).

43 Li, Q. *et al.* An objective metric of individual health and aging for population surveys. *Population health metrics* **20**, 11, doi:10.1186/s12963-022-00289-0 (2022).

44 López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. Hallmarks of aging: An expanding universe. *Cell* **186**, 243-278, doi:10.1016/j.cell.2022.11.001 (2023).

45 Nguyen, V. K. *et al.* Harmonized US National Health and Nutrition Examination Survey 1988-2018 for high throughput exposome-health discovery. *medRxiv*, doi:10.1101/2023.02.06.23284573 (2023).

46 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 medicine* **12**, e1001779 (2015).

47 Zhao, Y., Hu, Y., Smith, J. P., Strauss, J. & Yang, G. Cohort profile: the China health and retirement longitudinal study (CHARLS). *International journal of epidemiology* **43**, 61-68 (2014).

48 Gu, D., Feng, Q., Chen, H. & Zeng, Y. in *Encyclopedia of gerontology and population aging* 957-970 (Springer, 2022).

49 Pu, Y. *et al.* Gut microbial features and circulating metabolomic signatures of frailty in older adults. *Nat Aging* **4**, 1249-1262, doi:10.1038/s43587-024-00678-0 (2024).

50 Kang, L., Chen, W., Petrick, N. A. & Gallas, B. D. Comparing two correlated C indices with right-censored survival outcome: a one-shot nonparametric approach. *Stat Med* **34**, 685-703, doi:10.1002/sim.6370 (2015).

51 Uno, H., Cai, T., Pencina, M. J., D'Agostino, R. B. & Wei, L. J. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* **30**, 1105-1117, doi:10.1002/sim.4154 (2011).

Table 1. Associations of DISCO with mortality across five cohorts.

Cohorts	Follow-up years	Events/N	HR (95% CI)	DISCO C-index	DM C-index
NHANES	10.1 (3.8-13.1)	3304/27059	1.71 (1.68-1.75)	0.767	0.766
UKB					
Clinical	13.8 (13.0-14.5)	42770/487517	1.61 (1.60-1.62)	0.670	0.655
Proteomics		5117/47859	2.41 (2.35-2.47)	0.740	0.697
ClinicalProt		5022/47173	2.38 (2.33-2.44)	0.741	0.677
Validations					
CHARLS	8.0 (8.0-8.0)	1096/11601	1.26 (1.21-1.30)	0.689	0.670
CLHLS	3.3 (2.2-3.5)	828/2213	1.27 (1.21-1.33)	0.651	0.605
RuLAS					
Wave 2	4.0 (4.0-4.0)	190/1821	1.33 (1.21-1.46)	0.650	0.604
Wave 4	2.0 (2.0-2.0)	124/2120	1.45 (1.30-1.62)	0.678	0.662

Follow-up years were shown with median (interquartile). HR (95% CI) was estimated adjusting for age and sex, while C-index was calculated based on crude model (mortality ~ DISCO/DM). HR: hazard ratio; CI: confidence interval; DISCO: distance of covariance; DM: Mahalanobis distance;

Figures

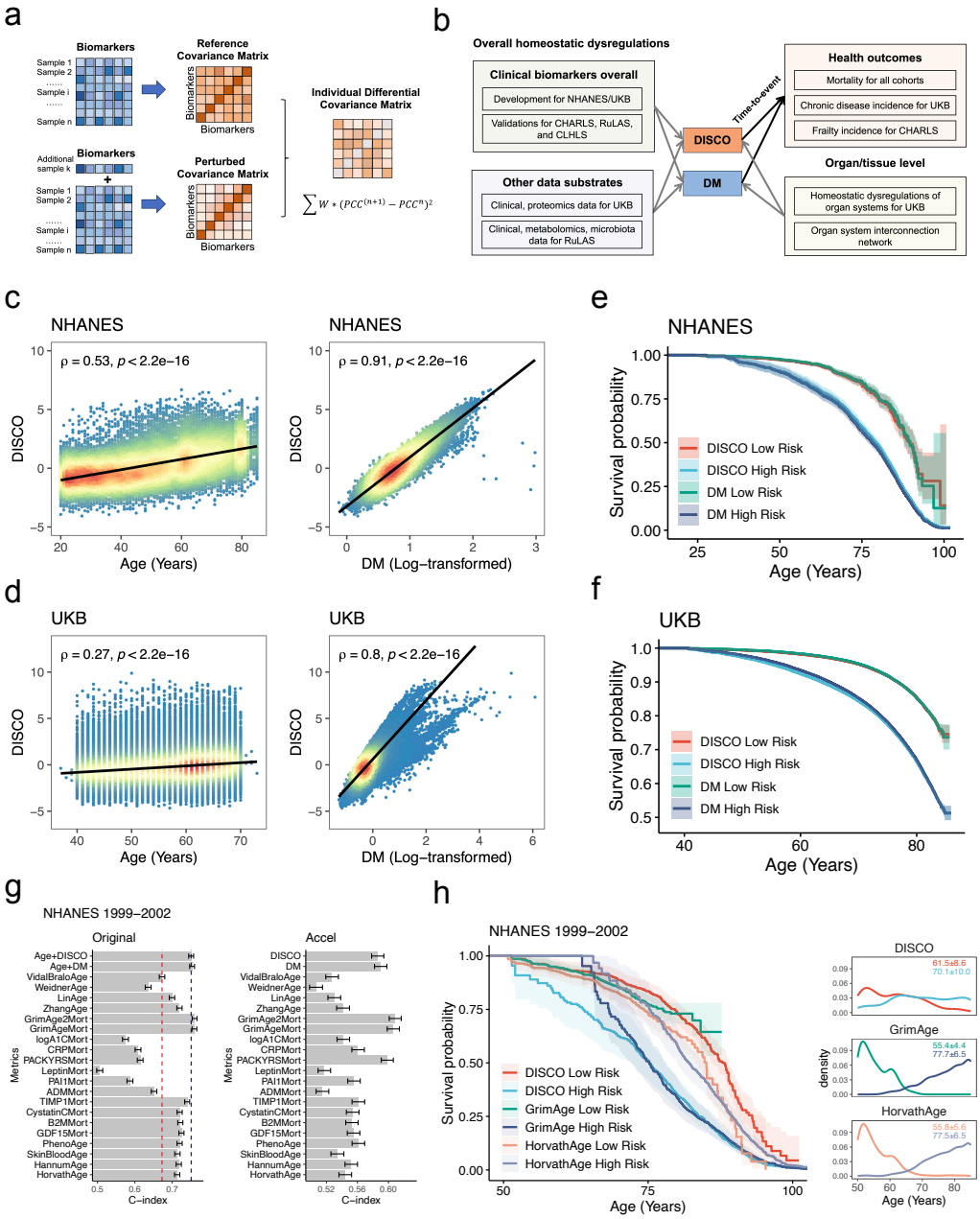
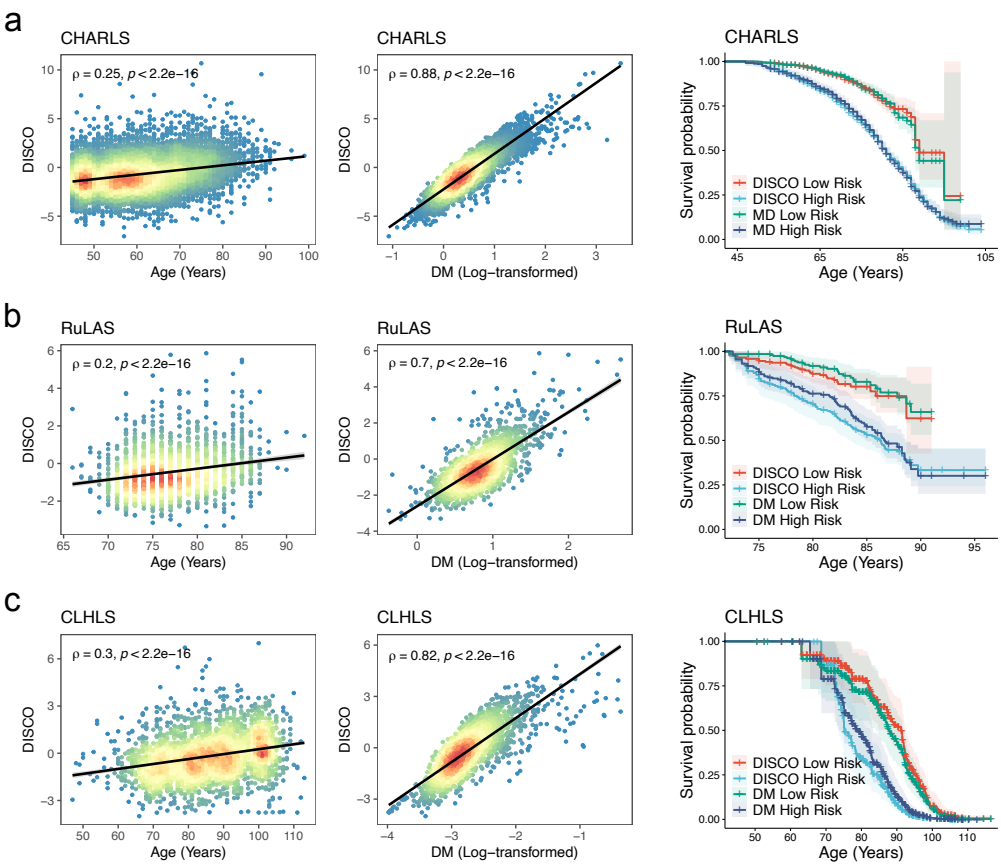


Figure 1. The DISCO framework and its application to clinical biomarkers. (**a**) Conceptual basis of distance of covariance (DISCO): The phenotypic covariance matrix changes slightly with an additional sample (e.g, $sample_k$), highlighting DISCO as an individual-level metric quantifying aggregated covariance difference. (**b**) The workflow of this study, which proposed the DISCO and applied it on various datasets to investigate the homeostatic dysregulations and interconnections of organs. Scatter plots (**c**) of DISCO versus age (left) and Mahalanobis distance (DM; right) in the NHANES cohort, with Spearman's correlation coefficients (ρ) and p values. Analogous plots (**d**) for the UK Biobank (UKB) cohort. Kaplan-Meier curves (**e, f**) compared survival probabilities for individuals categorized by low/high DISCO and DM risks

(first and forth quartiles) in NHANES and UKB, respectively, using chronological age as the time-scale axis. Bar plots (**g**) compared the discriminative accuracy (C-index) of DISCO against epigenetic clocks, epigenetic-derived markers, and DM for predicting mortality in NHANES. C-index of DISCO and age were marked with navy blue and red dotted line in the left panel, respectively. The right panel presented the C-index of age-adjusted acceleration (Accel), a residual-adjusted measure. Kaplan-Meier curves (**h**) compared survival probabilities for individuals categorized by low/high DISCO, GrimAge, and HorvathAge risks (first and forth quartiles) in NHANES and UKB, respectively. Right panels: Age distributions (mean ± s.d.) per risk group. Low mortality at younger ages in high-risk GrimAge and HorvathAge clocks in the Kaplan-Meier curves is likely attributable to the small sample at those ages in those groups.



Extended Figure 1. Validation of DISCO across three distinct cohorts of older adults. Scatter plots examined associations between the distance of covariance (DISCO), age, and Mahalanobis distance (DM) in CHARLS (**a**), RuLAS (**b**) and CLHLS (**c**), with each plot displaying Spearman correlation coefficients (rho) and p-values. Kaplan-Meier survival curves (right panels) compare mortality outcomes for participants categorized by quartile-based risk (the first and fourth quartiles) of DISCO and DM with the age as time-scale axis.

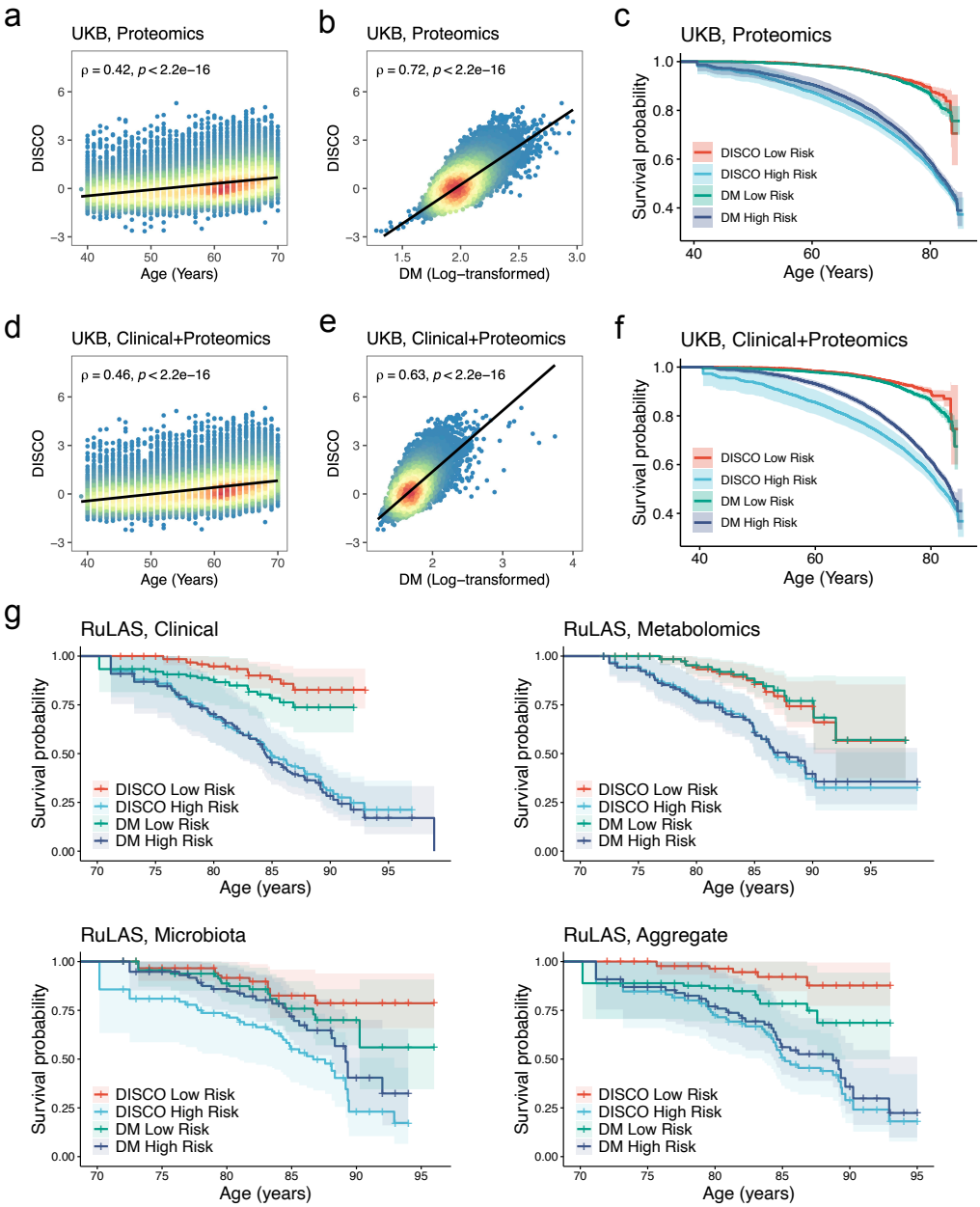
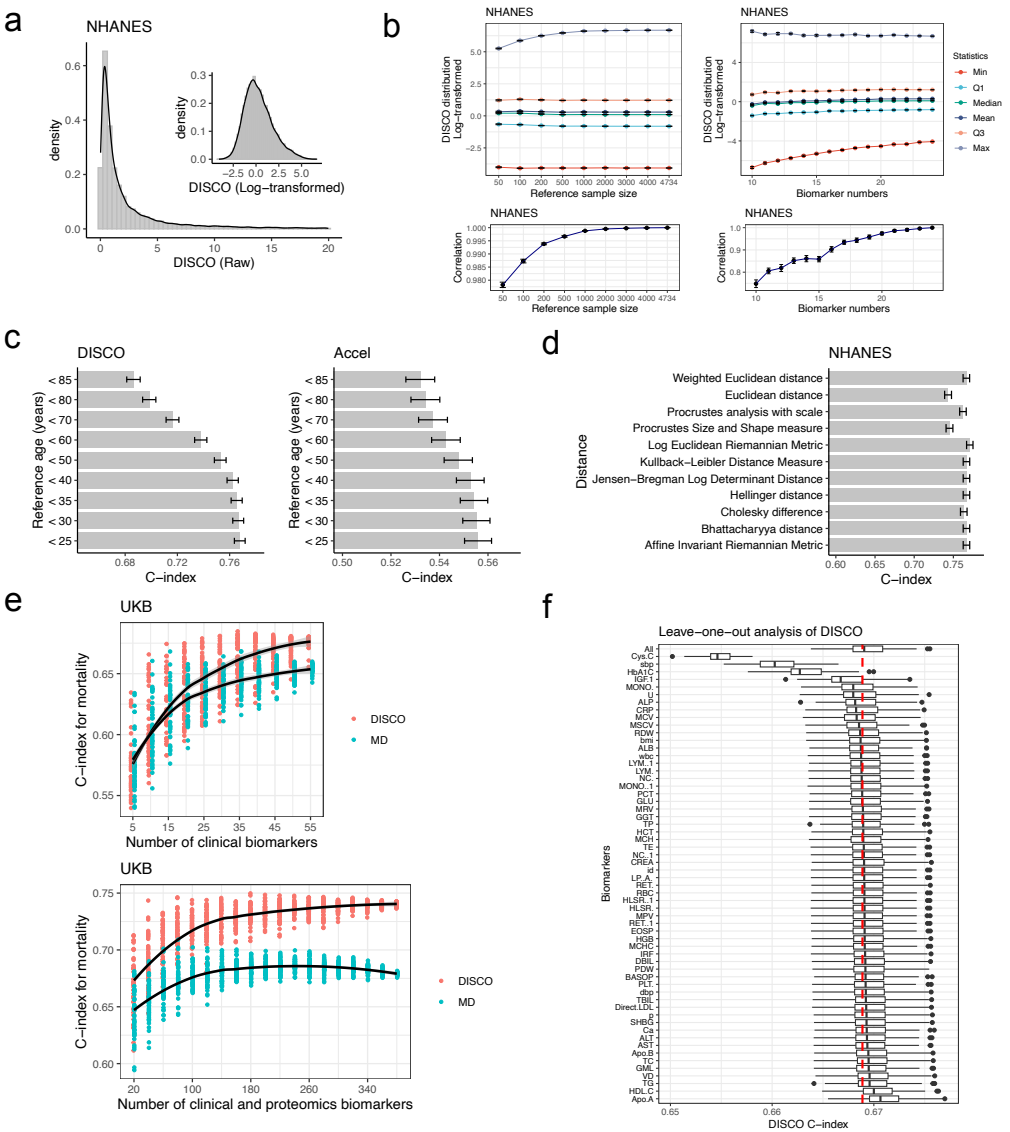


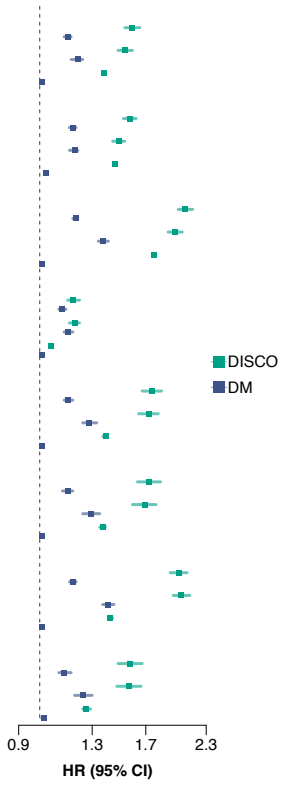
Figure 2. DISCO applications across diverse omics biomarker domains. The scatter plots of DISCO correlations with age and DM in UKB using proteomics (**a, b**) and union sets of proteomics and clinical biomarkers (**d, e**) in the UKB. Spearman correlation coefficients and p values were presented. The Kaplan-Meier curves (**c, f**) of the individuals stratified by the DISCO and DM low and high risks (the first and fourth quartiles) with age as the time-scale axis. Kaplan-Meier survival curves (**g**) stratified by DISCO quartiles from clinical, metabolomic, gut microbiota, and aggregate DISCO scores in the RuLAS cohort.



Extended Figure 2. Robustness analysis of DISCO. Density distribution plots (**a**) of raw and log-transformed DISCO values. The DISCO distribution (**b**) across different reference (young adult group) size and biomarker numbers in NHANES. The Spearman correlation of DISCO values for the 50 replicates at each parameter setting were evaluated. Bar plots (**c**) evaluated C-Index performance using reference datasets with different ages in the NHANES. Bar plots (**d**) evaluated C-Index performance using varying distance metrics to calculate DISCO in NHANES. Scatter plots (**e**) show C-Index distribution in predicting mortality through DISCO and DM with randomly selected clinical and proteomics biomarkers in the UKB (50 replicates per biomarker number). Boxplots (**f**) evaluated the stability of DISCO C- index in leave-one-out analysis, in which each biomarker was excluded one-at-a-time (50 replicates per biomarker).

a

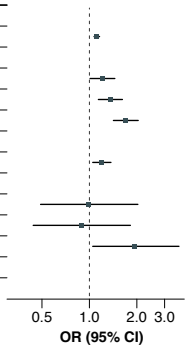
Diseases/Metrics	Events (%)	DISCO HR	DISCO C–index	DM HR	DM C–index
Obesity					
ClinicalProt	3752 (8.20%)	1.59	0.615	1.15	0.571
Prot	3812 (8.21%)	1.53	0.606	1.21	0.574
Clinical	38241 (8.06%)	1.38	0.616	1.01	0.594
Hypertension					
ClinicalProt	6799 (19.93%)	1.57	0.642	1.18	0.590
Prot	6910 (19.95%)	1.49	0.627	1.19	0.593
Clinical	68269 (19.04%)	1.46	0.655	1.03	0.615
Diabetes					
ClinicalProt	3252 (7.10%)	2.07	0.721	1.20	0.669
Prot	3315 (7.14%)	1.97	0.704	1.37	0.661
Clinical	32219 (6.79%)	1.77	0.746	1.01	0.709
Cancer					
ClinicalProt	7354 (17.18%)	1.18	0.592	1.12	0.566
Prot	7479 (17.22%)	1.19	0.590	1.15	0.573
Clinical	75784 (17.09%)	1.06	0.552	1.01	0.542
MI					
ClinicalProt	1811 (3.94%)	1.75	0.684	1.15	0.628
Prot	1841 (3.95%)	1.73	0.676	1.28	0.632
Clinical	17500 (3.68%)	1.39	0.645	1.01	0.621
Stroke					
ClinicalProt	1240 (2.67%)	1.73	0.681	1.15	0.621
Prot	1256 (2.67%)	1.69	0.672	1.29	0.631
Clinical	11895 (2.48%)	1.37	0.648	1.01	0.624
COPD					
ClinicalProt	2040 (4.43%)	2.01	0.739	1.18	0.660
Prot	2081 (4.45%)	2.03	0.738	1.41	0.682
Clinical	19151 (4.01%)	1.42	0.667	1.01	0.655
Dementia					
ClinicalProt	1279 (2.71%)	1.57	0.730	1.13	0.643
Prot	1297 (2.71%)	1.56	0.725	1.24	0.659
Clinical	9403 (1.93%)	1.26	0.644	1.02	0.630



b

CHARLS, Frailty incidence

Category	Events (%)	OR (95% CI)	P value
2011 to 2015			
All	1427 (15.55%)	1.11 (1.07–1.16)	<0.001
Q1	266 (11.41%)	1 (Ref)	
Q2	328 (13.74%)	1.21 (1.01–1.45)	0.035
Q3	378 (16.20%)	1.36 (1.14–1.62)	0.001
Q4	455 (21.41%)	1.70 (1.42–2.04)	<0.001
P for trend			<0.001
2015 to 2018			
All	297 (8.48%)	1.20 (1.05–1.37)	0.007
Q1	48 (5.34%)	1 (Ref)	
Q2	63 (7.14%)	0.99 (0.49–2.03)	0.980
Q3	72 (7.98%)	0.89 (0.44–1.82)	0.747
Q4	114 (13.92%)	1.92 (1.05–3.70)	0.042
P for trend			0.036



c

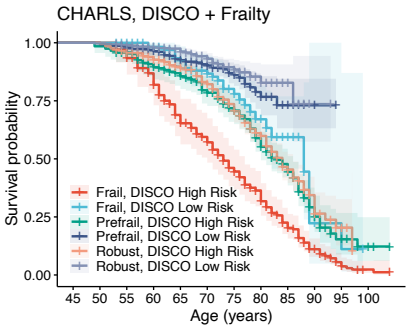
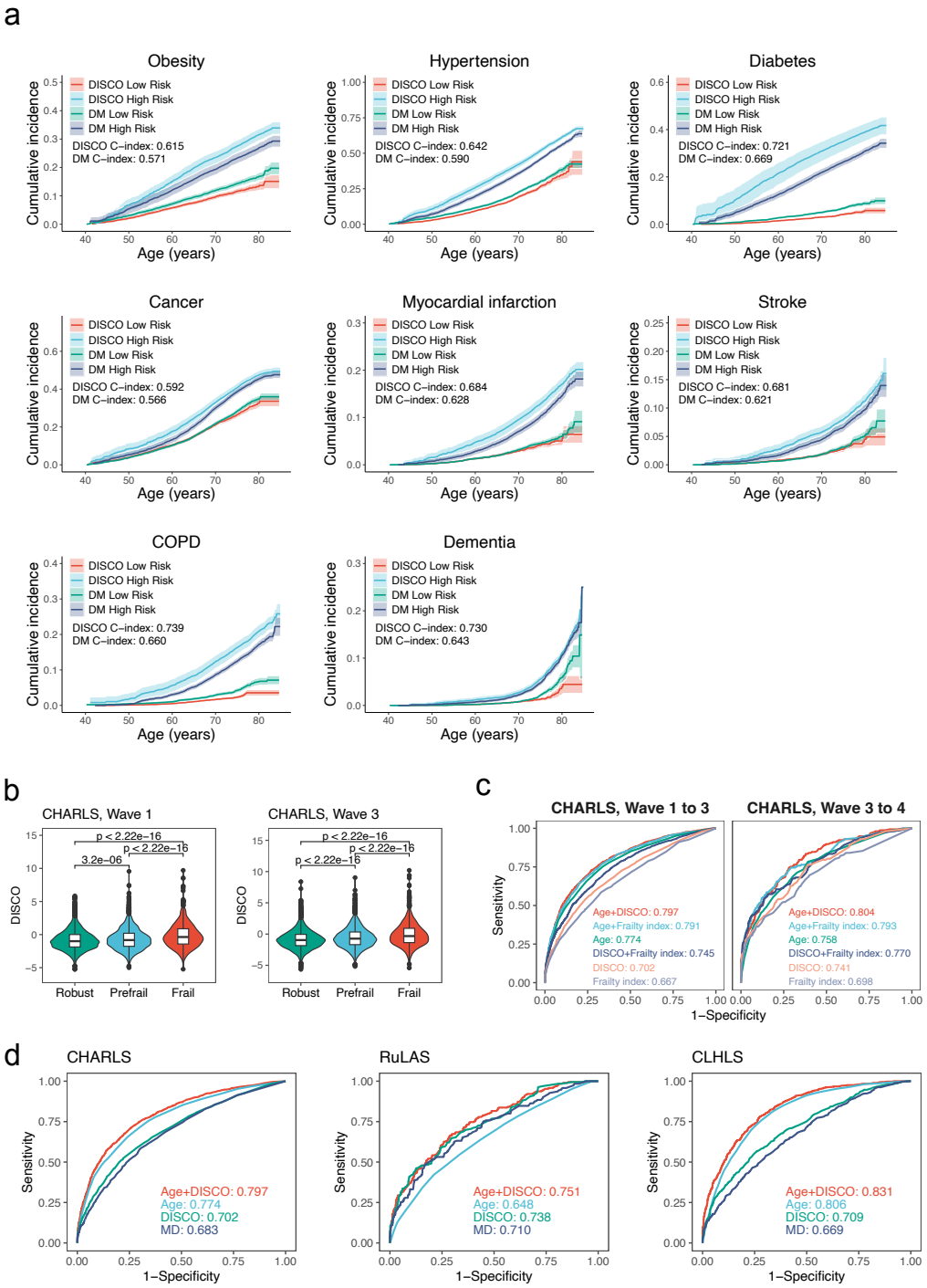


Figure 3. Assessment of DISCO's predictive capacity for incidence of chronic disease, frailty and mortality. Forest plot (**a**) presents hazard ratios (HRs, 95% confidence interval [CI]) for DISCO-associated chronic diseases in the UKB cohort (14-year follow-up). HRs (95% CI) were adjusted for age and sex (Model 2). C-index of Model 1 (unadjusted) was shown. DISCO and Mahalanobis distance (DM) were derived from proteomics (Prot), clinical biomarkers, and their combined set (ClinicalProt). Forest plot (**b**) of odds ratios (ORs, 95% CI) for incident frailty (frailty index ≥0.25, assessed via 32 functional and disease-related items) in the CHARLS cohort, validated across discovery (2011–2015) and validation (2015–2018) waves. Kaplan-Meier survival curves (**c**) stratified participants by DISCO risk quartiles and

frailty status. Model 1: Disease ~ DISCO/DM; Model 2: Disease ~ DISCO/DM + age + sex. MI: myocardial infarction; COPD: chronic obstructive pulmonary disease.



Extended figure 3. Predictive performance of DISCO for incident chronic disease and frailty. Kaplan-Meier survival curves (**a**) compared the cumulative incidence of chronic diseases (e.g., myocardial infarction [MI], chronic obstructive pulmonary disease [COPD]) in UKB. Individuals were categorized by low (first quartile) and high (fourth quartile) DISCO and

Mahalanobis distance (DM) risk groups, with the C-index was shown for each disease. Violin plots (**b**) of DISCO across frailty status which was assessed through frailty index (Robust: frailty index < 0.1; Prefrail: 0.1<= frailty index<0.25; Frail: frailty index >= 0.25). Differences were analyzed using the Wilcoxon tests. Receiver operating characteristic (ROC) curves (**c**) comparing discriminative accuracy of DISCO, frailty index, and age for 4-year (2011 to 2015) and 3-year (2015 to 2018) mortality prediction in CHARLS. The area under the curve (AUC) was shown. (**d**) ROC curves evaluated the discriminative accuracy of DISCO, DM, and age for predicting mortality in CHARLS (2011 to 2020), RuLAS and CLHLS (2014 to 2018). The area under the curve (AUC) was shown for each metric. DISCO: distance of covariance; DM: Mahalanobis distance; ROC: receiver operating characteristic curve; AUC: area under curve.

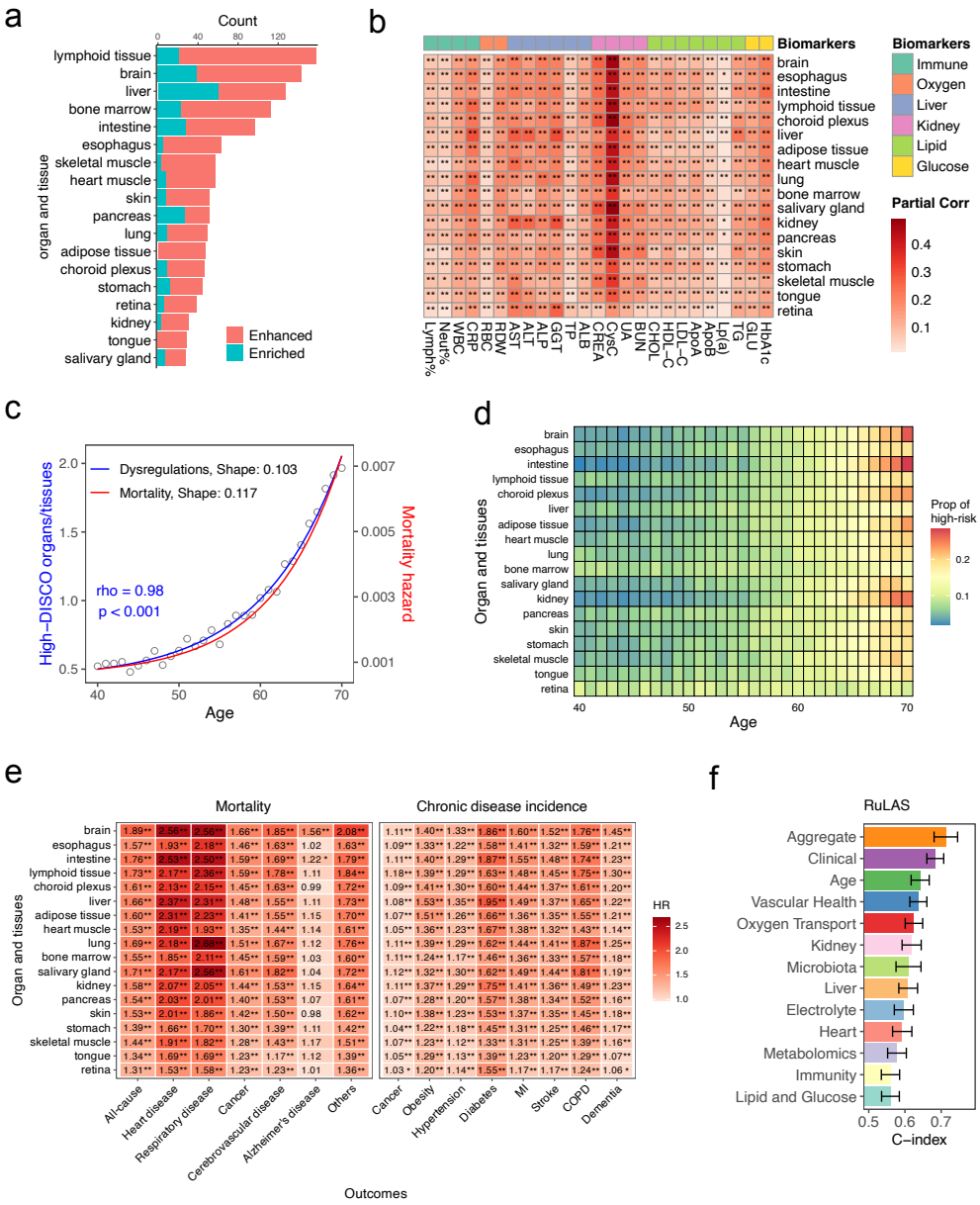
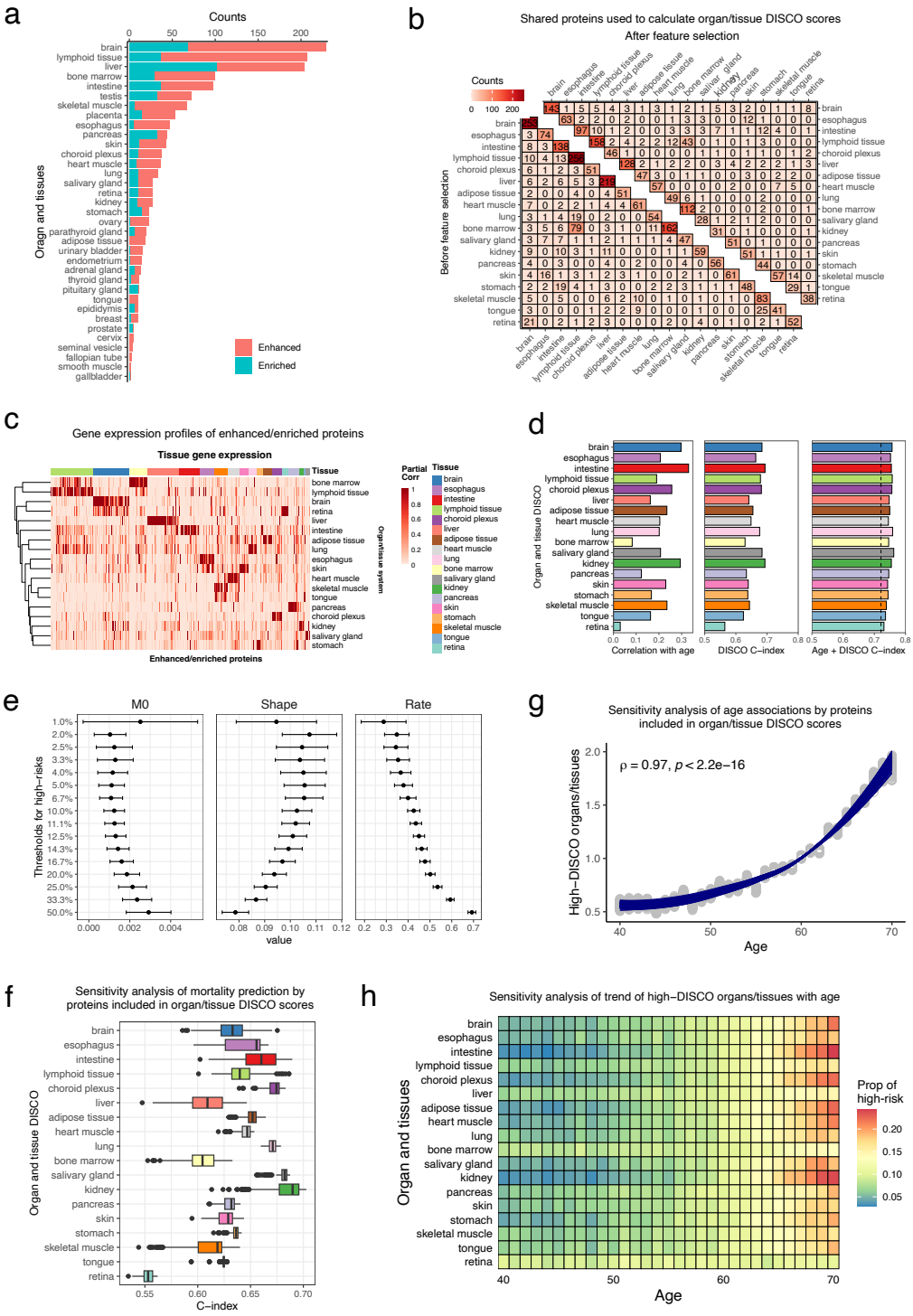


Figure 4. Multi-organ homeostatic dysregulations profiling by DISCO. Bar plots (**a**) of organ/tissue-specific protein counts (enhanced: ≥fourfold mRNA in one tissue vs. average; enriched: ≥ fourfold vs. all others) in UKB. A heatmap (**b**) visualizes the partial correlations (r) between DISCO values (18 organs/tissues) and routine aging biomarkers, adjusting for age and sex. (**c**) The ratio of observed/expected high-risk (top 10% DISCO scores) organs and tissues (black circles) increases exponentially with age; the blue line shows a fitted Gompertz-Makeham model. For comparison, the red line shows a Gompertz-Makeham fit to mortality. A heatmap (**d**) showed age-stratified high-risk organ prevalence (DISCO top 10%). Heatmaps (**e**) present hazard ratios (HRs, 95% confidence interval [CI]) for all-cause and cause-specific mortality, and chronic diseases. The statistical significances were shown with asterisks (**: $p<0.01$; *: $p<0.05$). MI: myocardial infarction; COPD: chronic obstructive pulmonary disease. Bar plot (**f**) displays the C-index for 2-year mortality prediction across organ systems defined through clinical, metabolomics and microbiota data in RuLAS. Standard errors (SEs) are shown as error bars.



Extended figure 4. Organ-specific protein signatures and homeostatic dysregulations. (**a**) Bar plots of enhanced and enriched proteins across organ and tissue in original UKB Olink proteomics panel (enhanced: ≥fourfold mRNA in one tissue vs. average; enriched: ≥ fourfold vs. all others). Matrix heatmaps (**b**) show the number of shared proteins between pairwise organ/tissue groups before (left lower triangle) and after (left upper triangle) feature selection. (**c**) The gene expression (scaled TPM values) profiles of enhanced and enriched proteins across

organs and tissues, obtained from Human Protein Atlas website, are shown in the heatmap. Bar plots (**d**) show Pearson correlations between organ-specific DISCO and age, C-index of x-year mortality prediction for DISCO alone, and C-index for DISCO + age for 18 organs and tissues. (**e**) Parameters of the Gompertz-Makeham model (M_0, Shape, Rate) were estimated using a set of quantiles of DISCO to define high-risk homeostatic dysregulations across age. Panels **f-h** show the results of sensitivity analysis using randomly selected protein (40 proteins per organ, 200 replicates). The C-index and multi-organ homeostatic dysregulation trends across age were evaluated.

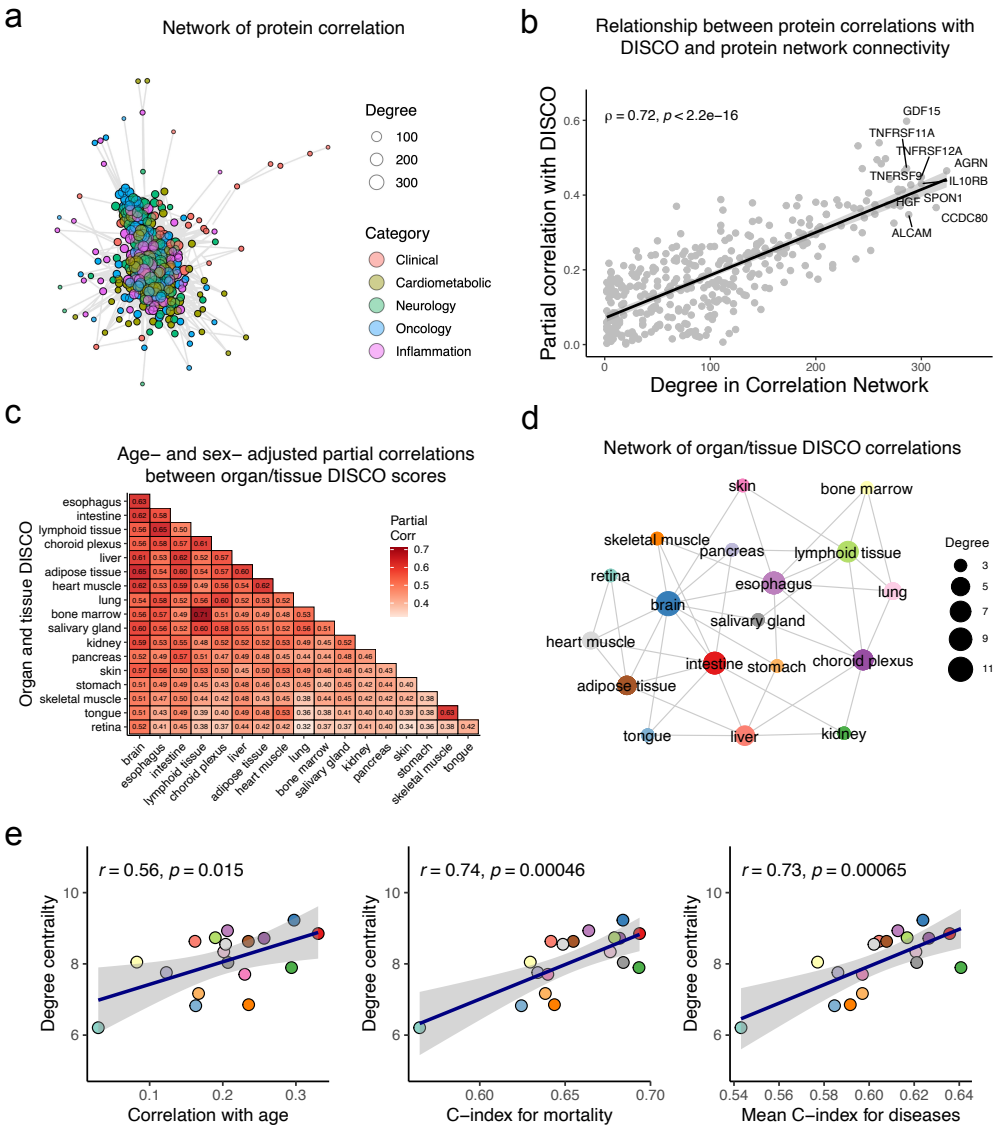
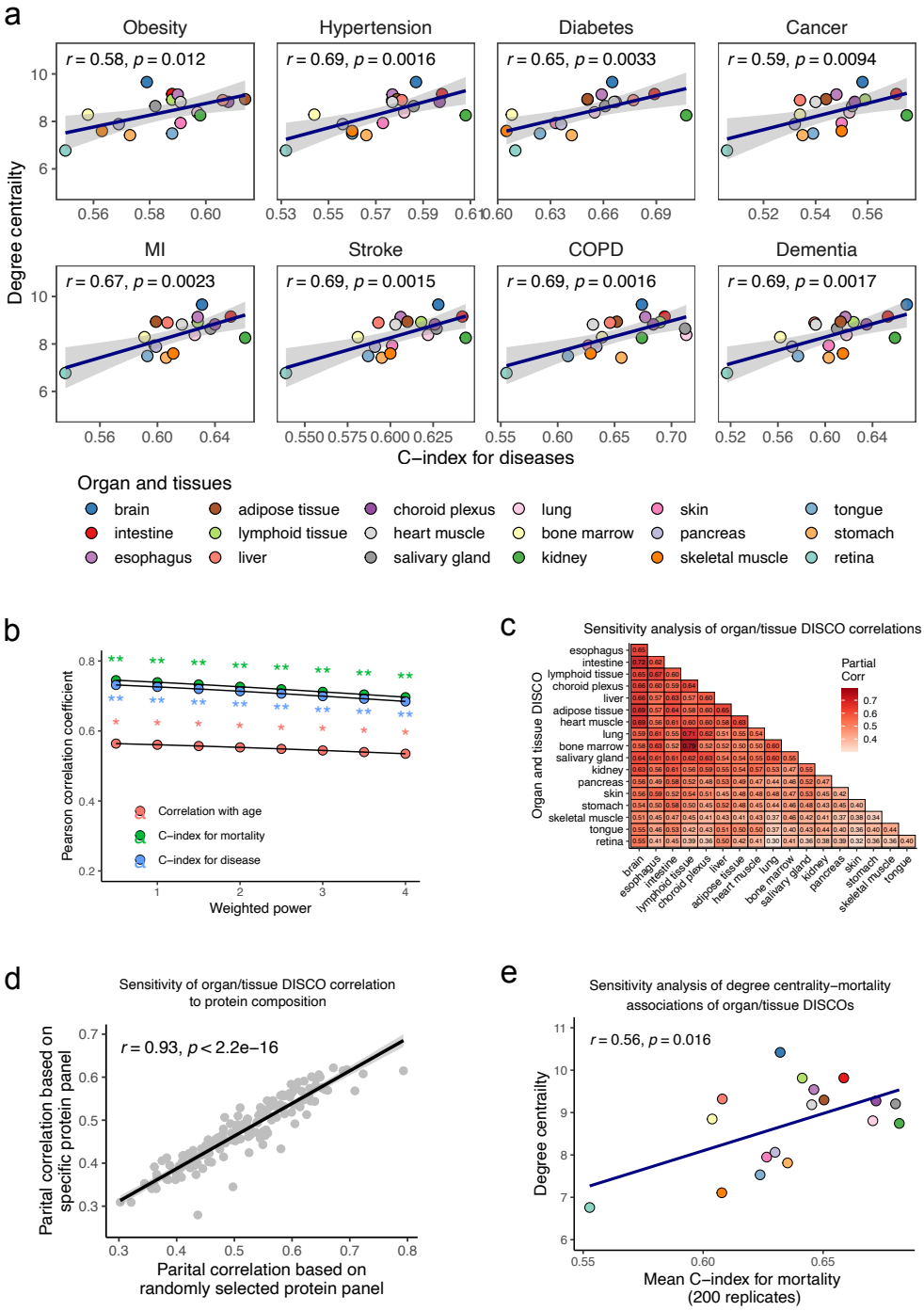


Figure 5. Multi-organ homeostatic dysregulation revealed by network centrality analysis. (**a**) A graph diagram of clinical and proteomic biomarkers (UK Biobank data), where nodes represent individual biomarkers and edges indicate pairwise correlations between them. A scatter plot (**b**) displays biomarker network degree versus DISCO partial correlations (x-axis: network degree; y-axis: partial correlation with DISCO, adjusted for age and sex). A heatmap

(**c**) shows the partial correlations between DISCO values of 18 organs/tissues, adjusted for age and sex. A K-neighbor (K=3) network (**d**) visualizes strong inter-organ/tissue dysregulation links among organs/tissues. € Scatter plots show relationships between organ/tissue degree centrality (sum of correlations with all other nodes) and three key dysregulation metrics, with node colors matching panel D. Significance: correlation coefficients and *p*-values (Pearson) are shown.



Extended figure 5. Sensitivity analysis of multi-organ homeostatic dysregulation. Scatter plots (**a**) show degree centrality of organs/tissues versus C-index for predicting chronic disease incidence (Pearson's r and p values) in UKB. Sensitivity analysis (**b**) of the weighting power parameter used to define degree centrality on the correlations in Fig. 5c. A heatmap (**c**) depicts age-/sex-adjusted partial correlations between organ-specific DISCO values, averaged across 200 replicates of randomly selected protein subsets (40 proteins per organ). The relationships (**d**) of pairwise inter-organ/tissue DISCO partial correlations (y-axis: raw proteomics data adjusted for age/sex; x-axis: 200 random protein subsets). (**e**) Sensitivity analysis (40 proteins per system, 200 replicates) of the association between organ/tissue degree centrality and mortality prediction C-index using DISCO derived from random protein subsets, with node colors matching panel A. Statistical significance is indicated by Pearson correlation coefficients (r) and p values.