From Psychological Distress to Multi-System Morbidity: A Transdiagnostic Internalizing Factor Predicts Chronic Disease Onset

Ang Li

al@ibp.ac.cn

State Key Laboratory of Cognitive Science and Mental Health, Institute of Biophysics, Chinese Academy of Sciences https://orcid.org/0000-0002-2309-9811

Yini He

State Key Lab of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.

Tian Gao

State Key Laboratory of Cognitive Science and Mental Health, Institute of Biophysics, Chinese Academy of Sciences

Xiya Liu

State Key Lab of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.

Jiahong Wang

Health Management Institute, The Second Medical Center & National Clinical Research Center for Geriatric Diseases, Chinese PLA General Hospital

Junxing Xian

Tianjin University

Yingjie Peng

State Key Lab of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.

Fengqian Xiao

State Key Laboratory of Cognitive Science and Mental Health, Institute of Biophysics, Chinese Academy of Sciences

Yan Yan

Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518066, China. https://orcid.org/0000-0002-6344-136X

Lei Wang

Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences https://orcid.org/0000-0002-7033-9806

Kaixin Li
Institute of Biophysics, Chinese Academy of Sciences
Changsheng Dong
State Key Lab of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.
Shangzheng Huang
State Key Lab of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences
Xiaohan Tian
State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100091, China.
Chaoyue Ding
School of Artificial Intelligence, University of Chinese Academy of Sciences.
Qi Li
Capital Normal University
Yuheng Huang
College of education, Capital Normal University
Bing Liu
Beijing Normal University https://orcid.org/0000-0003-2029-5187
Jing Xiao
Beijing Key Laboratory of Learning and Cognition, School of Psychology, Capital Normal University, Beijing 100048, China.

Article

Keywords:

Posted Date: May 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9436157/v1

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

Abstract

Physical health and mental well-being are bidirectionally linked. While the impact of somatic health on mental well-being is extensively investigated, the reciprocal “mind-to-body” trajectory—how psychological distress contributes to multi-system physical disease—remains underexplored, particularly in general populations. Here, we systematically investigated the biological pathways linking internalizing psychopathology to chronic disease onset by comparing the relative contributions of central (brain structural) and peripheral (multi-system biomarker) mediators. Using 14-year longitudinal data from 158,435 UK Biobank participants, we derived a transdiagnostic internalizing factor that prospectively predicted the onset of 28 chronic diseases across six physiological systems, with the highest-risk individuals experiencing disease onset up to 3.1 years earlier. Comprehensive phenotyping across 363 biomarkers revealed that peripheral biological dysregulation—particularly in the musculoskeletal and cardiovascular systems—exhibited associations with internalizing symptoms comparable to or stronger than brain structural alterations. Mediation analyses indicated that brain structure primarily mediated risks for neurological and cardiometabolic conditions, whereas peripheral biomarkers served as broad mediating pathways for a wider range of somatic outcomes, involving inflammation, metabolic dysregulation, and cardiovascular dysfunction. These findings reframe internalizing psychopathology as a systemic biological state, suggesting that integrating peripheral biological monitoring with dimensional psychopathology assessment could enhance risk stratification for chronic multimorbidity.

Introduction

No health without mental health[1]. Extensive prospective research has documented the “body-to-mind” pathway: physical frailty, organ dysfunction, and metabolic dysregulation in middle-aged and older adults predict the onset of depression, anxiety, and sleep disorders, mediated by systemic inflammation and brain structural alterations[2–5]. However, the reverse pathway—how psychological distress drives future physical disease—remains incompletely understood. Emerging evidence suggests that internalizing symptoms (depression and anxiety) predict all-cause mortality and health outcomes[6,7], yet the biological mechanisms translating psychological vulnerability into organ-specific disease risk are largely unknown.

Understanding this “mind-to-body” pathway is particularly critical in the general population, where mechanisms likely differ from those observed in severe psychiatric disorders. In clinical populations (e.g., schizophrenia), associations between mental illness and physical disease are confounded by psychotropic medications and profound lifestyle disruptions[8,9]. In contrast, subclinical internalizing symptoms in community-dwelling adults may exert pathogenic effects through chronic, medication-independent physiological stress[10,11]. Several candidate pathways have been proposed. Chronic psychological distress activates the hypothalamic-pituitary-adrenal (HPA) axis, resulting in sustained hypercortisolemia that impairs glucose metabolism, promotes visceral adiposity, and increases cardiovascular risk[12,13]. Concurrently, it dysregulates the autonomic nervous system, reducing heart rate variability and elevating sympathetic tone, which drives hypertension and cardiac strain[14,15]. Internalizing

symptoms also induce chronic low-grade inflammation, characterized by elevated pro-inflammatory markers (e.g., C-reactive protein, leukocyte count, neutrophil count), which accelerate atherosclerosis, insulin resistance, and cellular senescence[16,17]. Critically, these pathways do not operate independently but interact across interconnected physiological networks (e.g., heart-brain-gut-liver axis)[18], amplifying disease risk through cross-system propagation.

Despite this mechanistic framework, the extent to which these biological alterations actually mediate disease onset—and which systems contribute most substantially—remains poorly quantified. Research on internalizing disorders has predominantly focused on brain alterations—cortical thinning, hippocampal atrophy, and limbic dysfunction—which characterize the neural substrates of psychopathology itself[19,20]. However, peripheral biological systems (musculoskeletal, cardiovascular, metabolic, immune, and respiratory) may be equally important mediating pathways when examining how psychological distress translates into physical disease. These peripheral systems have direct connections to target organs affected by chronic disease, yet their mediating roles have not been systematically quantified. This gap is particularly important in aging populations, where age-related peripheral decline may amplify the pathogenic effects of psychological distress.

Here, we addressed this gap using 14 years longitudinal data from 158,435 UK Biobank participants (mean age 53.9 years, 54.5% female). We derived a transdiagnostic internalizing factor capturing shared variance across depression, anxiety, neuroticism, and sleep disturbances—a continuous phenotype aligned with hierarchical dimensional models (e.g., HiTOP) that has proven more stable and predictive than categorical diagnoses[21,22]. We then examined prospective associations between this factor and 28 chronic diseases spanning six physiological systems, and quantified biological mediators across multiple domains: brain structure (grey matter volumes across 233 regions), cardiovascular function (cardiac structure and function via MRI), musculoskeletal health (grip strength, bone mineral density, muscle composition), immune activity (leukocyte subsets, C-reactive protein), metabolic homeostasis (lipid profiles, glucose metabolism), and respiratory function (spirometry parameters). Through systematic mediation analyses, we aimed to characterize which biological pathways most substantially translate internalizing psychopathology into chronic disease risk, with particular attention to peripheral systems that may operate independently of brain structural changes.

Results

Study sample

In total, the analytical sample comprised 158,435 individuals who had complete internalizing symptom scores and were free of chronic diseases within two years after baseline assessment. The average age was 53.9 years (SD = 8.0), and 54.5% were women (n = 86,306). Further details on the study design and sample sizes for each analysis are provided in Fig. 1 and Supplementary Table 1.

Transdiagnostic internalizing factor

Using bi-factor exploratory structural equation modeling (Bi-ESEM), we developed a transdiagnostic internalizing factor structure comprising 18 items that captured core dimensions of internalizing psychopathology, including anxiety and depression symptoms, neuroticism, and sleep disturbances (Fig. 2a, Supplementary methods and Supplementary Table 2). The general internalizing factor showed strong loadings across all 18 items (range: 0.28–0.74), indicating robust common variance. Three specific factors—nervousness, worry, and depression—captured residual variance after accounting for the general factor. The model showed good fit (comparative fit index [CFI], 0.97; Tucker-Lewis index [TLI], 0.96; root mean square error of approximation [RMSEA], 0.040 [95% CI, 0.040–0.041], all exceeding conventional thresholds for good fit; Supplementary Table 3). Longitudinal analyses across four time points spanning 16 years revealed high temporal stability for both the factor structure (correlation range, 0.96–0.99; false discovery rate *Ps* < 0.05; Fig. 2b) and individual factor scores (correlation range, 0.71–0.82; *Ps* < 0.05; Fig. 2c). Detailed factor loadings and model specifications are provided in Supplementary Tables 4–6.

Prospective associations between the transdiagnostic internalizing factor and chronic diseases

Each standard deviation increases in the internalizing factor predicted higher risks of 21 of 28 chronic diseases (HR = 1.05–1.48; Fig. 3a; Supplementary Table 7). The strongest associations were observed for irritable bowel syndrome (IBS, HR = 1.48, 95% CI = 1.42–1.55), Parkinson's disease (PD, HR = 1.24, 95% CI = 1.16–1.32), and respiratory conditions including chronic obstructive pulmonary disease (COPD, HR = 1.22, 95% CI = 1.16–1.27) and bronchiectasis (HR = 1.20, 95% CI = 1.11–1.30). Cardiometabolic diseases also showed substantial associations, including type 2 diabetes (T2D, HR = 1.17, 95% CI = 1.13–1.21) and obesity (HR = 1.15, 95% CI = 1.12–1.18). No significant associations were detected for any of the four cancer types examined (lung, colorectal, prostate, and breast cancer). Individuals with high internalizing scores developed 11 diseases 0.8–3.1 years earlier than those with low scores (Fig. 3b), particularly cirrhosis (3.1 years), dementia (2.6 years), and PD (2.3 years) (Supplementary Table 8). At the system level, the strongest associations were seen for respiratory diseases (HR = 1.20, 95% CI = 1.15–1.25) and neurological diseases (HR = 1.19, 95% CI = 1.16–1.23]; Fig. 3c, Supplementary Table 9). For cardiometabolic, immuno-inflammatory, musculoskeletal, and miscellaneous systems, risks increased by 10–11% per standard deviation. Cumulative incidence curves confirmed consistently higher disease rates in the high-score group across all systems during 14 years of follow-up (Fig. 3d). Given their particularly strong associations with the internalizing factor, IBS and cirrhosis were presented separately as conditions outside the five primary physiological systems, both showing substantially elevated cumulative incidence in high-score individuals.

Association between transdiagnostic internalizing factor and multiple systematic biological indicators

Among 233 regional grey matter volumes examined in 9,089 participants, nine regions showed significant associations with the internalizing factor after false discovery rate (FDR) correction. Larger volumes were observed in limbic and striatal structures, including the bilateral putamen and right caudate nuclei, whereas smaller volumes were identified in prefrontal and insular cortices (|β| range = 0.021–0.041; Fig. 4a; Supplementary Table 10). These effects remained stable across imaging follow-ups (Supplementary Table 11), suggesting a reproducible neuroanatomical signature of internalizing psychopathology. Moreover, the spatial pattern of grey matter associations with the internalizing factor strongly overlapped with a multimorbidity measure derived from the cumulative count of the 21 chronic conditions (r = 0.41, $P = 1.27 \times 10^{-10}$; Supplementary Fig. 1 and Supplementary Table 12), highlighting shared neural substrates for psychiatric vulnerability and disease burden.

Beyond the brain, robust associations were observed with multiple physiological systems (Fig. 4b, Supplementary Table 13). Musculoskeletal and cardiovascular markers demonstrated notable associations with the internalizing factor. The weight-to-muscle ratio showed the strongest positive association (β = 0.059, $P = 2.09 \times 10^{-10}$), whereas reduced hand grip strength (β = -0.053, $P = 4.05 \times 10^{-95}$), lower bone mineral density (β = -0.0402, $P = 8.45 \times 10^{-23}$), and decreased maximum right atrial volume (β = -0.041, $P = 3.95 \times 10^{-6}$) were associated with vulnerability across structural and functional domains. Additional cardiovascular parameters (e.g., left ventricular mass and stroke volume; |β| range = 0.022–0.039, $Ps < 3.95 \times 10^{-3}$) and musculoskeletal measures (vitamin D, thigh muscle volume, muscle fat infiltration, phosphate; |β| range = 0.011–0.032, $Ps < 3.78 \times 10^{-2}$) also demonstrated significant associations.

Other peripheral systems showed convergent evidence. In the metabolic domain, higher factor scores were linked to elevated triglycerides (β = 0.027, $P = 4.22 \times 10^{-15}$) and reduced HDL cholesterol (β = -0.017, $P = 2.04 \times 10^{-7}$) and apolipoprotein A (β = -0.012, $P = 6.34 \times 10^{-4}$). Immune dysregulation was reflected in higher neutrophil counts, neutrophil percentages, total white blood cell counts, and C-reactive protein levels (β range = 0.0096–0.025, $Ps < 6.74 \times 10^{-3}$). In the respiratory system, poorer lung function was indicated by reduced peak expiratory flow (β = -0.025, $P = 9.92 \times 10^{-27}$), forced expiratory volume in 1 second (β = -0.013, $P = 1.09 \times 10^{-8}$), and forced vital capacity (β = -0.010, $P = 3.63 \times 10^{-5}$). These patterns were consistent over time (Supplementary Table 14). Notably, associations between systemic physiological indicators and the multimorbidity measure showed similar patterns to those observed for the internalizing factor (Supplementary Table 15; Supplementary Fig. 2), suggesting that common biological pathways may link psychiatric vulnerability with chronic physical disease risk.

Biological indicators mediating the association between the transdiagnostic internalizing factor and chronic diseases

Having established robust associations between the internalizing factor and biological markers across neural and peripheral systems, we next examined whether these markers mediate the link between the factor and chronic disease risks. Mediation analyses revealed that biological systems contributed differentially to the association between the transdiagnostic internalizing factor and chronic diseases across the five systems (Fig. 5a-e, Supplementary Table 16). For cardiometabolic diseases, the mediation of the association involved multiple biological systems, with metabolic (indirect effect = 0.0053, $P < 0.001$) and cardiovascular markers (indirect effect = -0.0099, $P < 0.001$) served as primary mediators. For neurological diseases, the mediation primarily through musculoskeletal system (indirect effect = 0.0062, $P < 0.001$) and neurological system (indirect effect = 0.0061, $P = 0.027$). Respiratory and immuno-inflammatory diseases shared common mediating pathways through respiratory and immune systems, with additional system-specific contributions from musculoskeletal markers for respiratory diseases and metabolic markers for immuno-inflammatory diseases. For musculoskeletal diseases, the mediation primarily through cardiovascular (indirect effect = − 0.0043, $P = 5.14 \times 10^{-3}$) and musculoskeletal markers (indirect effect = 0.0016, $P = 0.048$). Given IBS's particularly robust association with the internalizing factor and its location outside the five core systems, we conducted a supplementary analysis focused on IBS (Fig. 5f). This analysis revealed that biomarkers from the cardiovascular system (indirect effect = 0.0102, $P = 2.72 \times 10^{-2}$) and respiratory system (indirect effect = 0.0016, $P = 2.27 \times 10^{-2}$) mediated this relationship.

Discussion

By integrating longitudinal clinical data, neuroimaging, and comprehensive physiological biomarkers in a large population-based cohort, we identified a temporally stable transdiagnostic internalizing factor in middle-aged and older adults that predicted 21 chronic diseases across multiple systems, with the strongest association observed for IBS (48% increased risk). The factor was also associated with an earlier age of onset for 11 chronic diseases. Our findings challenge the prevailing brain-centric paradigm in psychiatric research: while the internalizing factor correlated with grey matter alterations in emotion-regulation regions, peripheral biomarkers—particularly musculoskeletal, cardiovascular, respiratory, and immune markers—showed comparable and robust associations with the internalizing factor and mediated disease risks across a broader range of physiological systems. This work reveals that psychological distress manifests as a systemic biological state extending far beyond the brain, with profound implications for understanding and preventing chronic multimorbidity in individuals with internalizing disorders.

In middle-aged and older adults—the demographic bearing the highest burden of chronic multimorbidity —we identified a transdiagnostic internalizing factor that exhibited exceptional stability across 16 years (factor structure correlations: 0.96–0.99), establishing it as a reliable marker for long-term disease risk assessment. The bi-factor model captured shared variance across depression, anxiety, neuroticism, and sleep disturbances while preserving specificity for individual symptom domains, consistent with hierarchical models of psychopathology[23,24]. The general internalizing factor demonstrated strong

loadings across all 18 items (range: 0.28–0.74), alongside three specific factors that captured residual variance in nervousness, worry, and depression symptoms. Beyond structural validity, individual factor scores exhibited exceptional test-retest correlations (r = 0.71–0.82 across 16 years), indicating that the factor captures enduring trait-like vulnerability rather than fluctuating mood states. This long-term stability extends prior work supporting dimensional approaches to psychopathology[25,26], which has predominantly established such stability in younger populations[27], whereas our findings demonstrate that the transdiagnostic internalizing factor can serve as a reliable marker in aging populations facing peak chronic disease incidence.

The internalizing factor predicted elevated risk for 21 of 28 chronic diseases examined, with disease-specific patterns revealing differential vulnerability across physiological systems. The strongest association emerged for IBS (HR = 1.48), followed by Parkinson's disease (HR = 1.24) and COPD (HR = 1.22), conditions characterized by prominent stress-sensitivity and autonomic dysregulation. Consistent with prior research, internalizing disorders were closely linked to elevated risk and earlier onset of neurodegenerative diseases[28,29]: higher scores raised Parkinson’s disease risk by 24% and dementia by 18%, with onset 2.3 and 2.6 years earlier, respectively. The increased burden of cardiometabolic disorders may partly reflect adverse health behaviors and metabolic dysregulation previously associated with internalizing symptoms[30,31], with higher risks for type 2 diabetes (HR = 1.17) and obesity (HR = 1.15), accompanied by earlier onset of 1.7 and 1.3 years. Notably, cirrhosis showed the most pronounced shift in age at onset (3.1 years earlier), highlighting how chronic stress may accelerate progressive organ damage. Heightened risks for respiratory and immune-inflammatory conditions, including COPD (HR = 1.22), bronchiectasis (HR = 1.20), and rheumatoid arthritis (HR = 1.14), though without significant onset shifts, may plausibly involve chronic sympathetic overactivation and systemic stress responses[32,33]. Musculoskeletal disorders showed relatively modest increases[34], such as an 11% higher risk of osteoarthritis, accompanied by an earlier onset of 1.1 years. Collectively, these differential patterns suggest that the internalizing factor confers vulnerability to chronic diseases through multiple pathophysiological mechanisms, with the magnitude and timing of effects varying by disease system.

Our findings highlight the critical importance of peripheral biomarkers in understanding how internalizing psychopathology contributes to chronic disease risk. Although a few prior studies have suggested that internalizing disorders are linked to multi-system health outcomes[35,36], our systematic investigation across brain and peripheral systems reveals that psychological distress manifests as widespread biological alterations extending far beyond neural circuits. Among 233 brain regions examined, 9 showed significant associations after FDR correction ($|\beta|$ range = 0.021–0.041), including the bilateral orbitofrontal cortex, insula, and striatal structures—regions consistent with established emotion-regulation circuits[37,38]. In contrast, musculoskeletal markers such as weight-to-muscle ratio[39] (β = 0.059, $P = 2.09 \times 10^{-10}$) and grip strength[40] (β = − 0.053, $P = 4.05 \times 10^{-95}$) exhibited stronger associations ($|\beta|$mean = 0.032) than grey matter volumes ($|\beta|$mean = 0.026). A comparable pattern was observed for cardiovascular markers ($|\beta|$mean = 0.033), which reflected features of myocardial fibrosis, including reductions in left ventricular mass and end-diastolic volume[41].

System-specific mediation analyses reveal that peripheral biomarkers exert broader influences on chronic disease risk than brain structural measures. Brain structural metrics acted as mediators solely for neurological and cardiometabolic diseases, whereas musculoskeletal, respiratory, and immune biomarkers mediated associations across four distinct disease systems, demonstrating their more pervasive role in linking psychological distress to physical health outcomes. For cardiometabolic diseases, biomarkers from all six physiological systems contributed mediating effects, underscoring the multi-system biological embedding of these conditions. These patterns align with established evidence underscoring the intricate interplay between internalizing disorders and dysregulation of the HPA axis and sympathetic nervous system (SNS)[12,42,43]. Aberrant HPA axis activity disrupts cortisol and growth hormone secretion, fostering chronic inflammation, skeletal muscle atrophy, and compromised neuroimmune function[44–46], thereby amplifying risks for immune-inflammatory, musculoskeletal, and neurological diseases. Concurrently, SNS hyperactivation exacerbates gastrointestinal dysmotility, muscle tension, tachycardia, and respiratory distress, elevating risks for musculoskeletal, cardiometabolic, and respiratory disorders[47–49]. The particularly strong associations with musculoskeletal markers—including grip strength, bone density, and muscle composition—may reflect chronic cortisol-mediated protein catabolism and age-accelerated sarcopenia, processes that independently predict mortality and functional decline[50]. Intriguingly, the negative cardiovascular mediation effect reveals a pattern diverging from classical heart disease pathophysiology. Whereas typical cardiovascular diseases (heart failure, cardiomyopathy) are characterized by pathological cardiac enlargement, internalizing symptoms were associated with diminished cardiac volumes—a pattern more consistent with accelerated cardiac aging rather than overt disease[51,52].

Several limitations affect interpretation. First, the assessment of anxiety and depression at baseline was limited to four items from the PHQ-4 scale, and continuous measurements of symptoms related to other common internalizing disorders were not included in the UK Biobank assessment. Future studies should incorporate broader dimensions to construct a more comprehensive transdiagnostic internalizing factor. Second, most multi-system biomarkers in our sample were assessed at single time points, precluding definitive causal inference regarding temporal relationships. While our 2-year lag period and exclusion of baseline disease strengthen causal inference, repeated biomarker measurements would enable examination of dynamic trajectories. Third, UK Biobank participants are generally healthier than the general population[53], which may attenuate true associations and limit generalizability; replication in populations with higher disease burden and greater diversity is essential. Fourth, several important biomarkers warrant inclusion in future studies, including brain functional connectivity features, specific cytokines, stress hormones, and microbiome profiles[54,55]. Future research should prioritize longitudinal biomarker data and validation in diverse populations with greater disease burden.

In conclusion, we identified a transdiagnostic internalizing factor in middle-aged and older adults that serves as a robust predictor of chronic disease burden across multiple physiological systems over 14 years. Critically, the association between psychological distress and chronic disease risk operates through biological pathways extending across both the brain and peripheral systems—including

musculoskeletal, cardiovascular, immune, metabolic, and respiratory markers—revealing that internalizing psychopathology manifests as a systemic biological state rather than an isolated neural dysfunction. These findings also have immediate translational implications for healthy aging. The combination of stable psychological markers and accessible peripheral biomarkers—such as grip strength (measurable with less than $100 equipment), inflammatory markers (standard blood panels), and basic metabolic parameters—could enable risk stratification in primary care settings. This cost-effective approach could help identify at-risk individuals early, offering an opportunity for timely interventions aimed at both mental and physical health, thereby preventing the onset of chronic multimorbidity.

Methods

Study Design and Population

This population-based retrospective cohort study utilized data from the UK Biobank, a large-scale biomedical database established in 2006. Between 2006 and 2010, the study recruited approximately 500,000 participants aged 37 to 73 years from the general population across 22 assessment centers in the United Kingdom[56]. Participants completed comprehensive baseline assessments including touchscreen questionnaires[57] covering psychosocial factors, lifestyle, and sociodemographic characteristics, along with physical measurements and biological sampling. Between 2014 and 2020, a subset of participants underwent detailed imaging assessments[58]. The UK Biobank study was approved by the North West Multicenter Research Ethics Committee (No.11/NW/0382), and written informed consent was obtained from all participants. This research was conducted using the UK Biobank resources under application number 85139, and all participants provided written informed consent.

Assessment of internalizing disorder

Internalizing disorder symptoms were assessed at baseline using three validated instruments administered on a touchscreen questionnaire. The Patient Health Questionnaire-4 (PHQ-4) evaluated depression and anxiety symptoms experienced over the previous two weeks[59]. Each item was scored from 0 (not at all) to 3 (nearly every day), with responses summed to create a total score. The Eysenck Personality Questionnaire-Revised Short Form (EPQ-N)[60] assessed neuroticism through 12 binary (yes/no) items. For each symptom, participants' response of "yes" was coded as 1 and "no" as 0, with higher total scores indicating greater levels of neuroticism[61] Sleep quality was evaluated using standardized questions covering seven domains: sleep duration, getting up in the morning, chronotype, daytime napping, insomnia, snoring, and daytime sleepiness[62]. The specific questions for each item and the distribution of scores across the population are provided in Supplementary Table 17 and Supplementary Fig. 3.

Disease Ascertainment and Follow-up

We identified incident chronic diseases through linkage to national health registries until April 1, 2022, using the UK Biobank's 'First occurrences' category (Category ID 1712). This category documents participants' diagnoses and their earliest occurrence dates, integrating data from three sources: general practice primary care data, inpatient hospital records, and death registry data[63]. To establish temporal precedence in our analysis of transdiagnostic internalizing factor's predictive value, we excluded participants diagnosed with chronic diseases prior to or within two years after the baseline assessment. Our study examined 28 chronic physical diseases spanning five major physiological systems and a miscellaneous system. These diseases were selected according to four key criteria: First, they exhibit high prevalence among middle-aged and elderly populations, as previously reported in the literature[64]. Second, there are well-established associations with severe adverse health outcomes[65]. Third, they show documented comorbidity patterns with internalizing disorders[35,66]. Finally, based on calculations in the UK Biobank cohort, we limited our analyses to chronic diseases whose onset typically occurs at age 45 or older. The neurological diseases included multiple sclerosis (MS), PD, dementia, and stroke. Cardiometabolic diseases encompassed T2D, heart failure (HF), ischemic heart disease (IHD), hypertensive diseases, atrial fibrillation (AF), and obesity. We examined respiratory diseases including COPD and bronchiectasis. Immuno-inflammatory diseases comprised chronic enteritis, psoriatic arthritis (PsA), and rheumatoid arthritis (RA). Musculoskeletal conditions included osteoarthritis (OA), spinal disorders, and osteoporosis (OP). Miscellaneous diseases encompassed IBS, cirrhosis, chronic kidney disease (CKD), anemia, glaucoma, and eczema or dermatitis. Finally, we included four types of cancer: lung, colorectal, prostate, and breast cancer. Detailed ICD-10 codes and onset age distributions of the 28 chronic diseases are available in Supplementary Tables 18,19.

Assessment of multisystem biological indicators

Between 2014 and 2020, UK Biobank collected MRI imaging data across multiple organs, including structural MRI scans of the brain from more than 40,000 participants. These scans were conducted using a Siemens Skyra 32-channel 3T scanner. Imaging employed a magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence with the following parameters: repetition time (TR) = 2000 ms, echo time (TE) = 2.01 ms, 208 sagittal slices, flip angle = 8°, field of view (FOV) = 208 × 256 × 256 mm, matrix size = 256 × 256, and slice thickness = 1.0 mm, with 1 mm isotropic resolution. Further details of the imaging protocol are available at http://www.fmrib.ox.ac.uk/ukbiobank/protocol/V4_23092014.pdf.

High-resolution T1-weighted structural MRI data were analyzed using voxel-based morphometry (VBM), implemented via the Computational Anatomy Toolbox (CAT12) integrated within the SPM12 framework[67]. We combined brain structural imaging data from 9,089 participants who had complete internalizing symptoms and covariate data, and who were free of chronic diseases both at baseline and within two years of follow-up. All scans were first aligned with a T1-weighted template, segmented into grey matter, white matter, and cerebrospinal fluid, and then normalized to the MNI standard space using the Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (DARTEL) algorithm[68]. Segmented grey matter maps were modulated to preserve volume information and smoothed using a

Gaussian kernel with a full width at half maximum (FWHM) of 4 mm. All images underwent visual inspection to exclude artifacts or major preprocessing errors. Subsequent analyses were restricted to voxels within a grey matter mask defined by averaging the smoothed grey matter maps across subjects and applying a threshold of 0.1 to exclude non-brain voxels. Then we integrated multiple brain atlases (Human Brainnetome Atlas[69] for cortical and subcortical regions, a probabilistic atlas[70] for cerebellar regions and hypothalamic atlas[71] for hypothalamus regions) to analyze grey matter volumes across 233 regions. Directly derived from the Human Brainnetome Atlas were 210 cortical regions (105 subregions per hemisphere) and 12 basal ganglia regions (bilateral ventral/dorsal caudate, globus pallidus, ventral pallidum, ventral/dorsal putamen). Additionally, 11 regions (bilateral amygdala, hippocampus, thalamus, hypothalamus, and left/right/vermis cerebellum) were generated by integrating subregions from their respective atlases.

Cardiovascular MRI was conducted using a 1.5T Siemens MAGNETOM Aera scanner (Syngo Platform VD13A, Siemens Healthcare, Germany) following the UK Biobank imaging protocol[72]. The protocol included short-axis, long-axis, and aortic cine imaging sequences, all acquired using balanced steady-state free precession (bSSFP). Short-axis cine sequences comprised a full stack covering both ventricles, allowing quantification of ventricular volumes and function. Long-axis cine images included horizontal, vertical, and left ventricular outflow tract (LVOT) views. Aortic cine imaging was performed in the transverse plane at the level of the pulmonary trunk to assess aortic compliance, with accompanying brachial blood pressure measurements for central pressure estimation.

Building upon a previously published framework[41], we implemented an automated analysis pipeline utilizing a Fully Convolutional Network (FCN) architecture[73] (detailed in Supplementary Methods) in 14,288 participants to extract 82 cardiac measures. These included 58 myocardial metrics (global and regional wall thickness, circumferential strain, radial strain, and longitudinal strain), 18 chamber function parameters (ventricular and atrial ejection fraction, stroke volume, maximum/minimum volumes, left ventricular mass, and cardiac output), and 6 aortic indicators (maximum area, minimum area, and distensibility for both ascending and descending aorta).

This study utilized data on 18 immune-inflammatory biomarkers collected during the UK Biobank baseline assessment (2006–2010). After merging with questionnaire and covariate data, a total of 129,415 participants with complete data were included in the analysis. CRP levels were obtained from the blood biochemistry panel (Category ID: 17518), measured using a high-sensitivity immunoturbidimetric assay on the Beckman Coulter AU5800 analyzer. Platelet counts, total white blood cell counts, and 10 leukocyte subtypes (expressed in 10^9 cells/L) were derived from the hematological assay panel (Category ID: 100081). In addition to analyzing individual markers, we calculated five composite immune-inflammatory indices that have been shown in prior studies to be associated with asthma, cardiovascular diseases, and chronic inflammatory conditions[74,75]. These indices were computed using the following formulas:

- Systemic Immune-Inflammation Index (SII):

$$SII = Platelet\ Count \times\ Neutrophil\ Count/Lymphocyte\ Count$$

,

- Platelet-to-Lymphocyte Ratio (PLR):

$$PLR = Platelet\ Count/Lymphocyte\ Count$$

,

- Neutrophil-to-Lymphocyte Ratio (NLR):

$$NLR = Neutrophil\ Count/Lymphocyte\ Count$$

,

- Eosinophil-to-Neutrophil Ratio (ENR):

$$ENR = Eosinophil\ Count/Neutrophil\ Count$$

,

- Eosinophil-to-Lymphocyte Ratio (ELR):

$$ELR = Eosinophil\ Count/Lymphocyte\ Count$$

.

Metabolic assessment included 9 serum markers (n = 92,177) capturing lipid profiles (apolipoproteins, cholesterol fractions) and glucose regulation (glucose, HbA1c). Musculoskeletal evaluation comprised 17 indicators from three modalities: blood biochemistry (e.g., calcium, vitamin D), physical anthropometrics (hand grip strength and bone mineral density), and abdominal MRI–derived body composition measures (sample sizes detailed in Supplementary Table 1). Respiratory function was assessed using spirometry parameters (n = 132,319), including forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1), peak expiratory flow (PEF), and FEV1/FVC ratio.

Covariates

Based on prior research—particularly studies leveraging UK Biobank data, we selected covariates known to be strongly associated with internalizing symptoms and the onset of chronic diseases[7,76], considering both sociodemographic and behavioral factors. Specifically, we included sex (Field ID 31), age (Field ID 21022), BMI (Field ID 21001), socioeconomic status (represented by the Townsend Deprivation Index, TDI, Field ID 22189), qualifications (Field ID 6138), smoking status (Field ID 20116), and alcohol intake frequency (Field ID 1558) as foundational covariates. Furthermore, in analyses related to brain structure, we added three additional covariates: head size (represented by volume scaling from T1 head images to

standard space, Field ID 25000), site information (Field ID 54), and scanner table position (Field ID 25759).

Statistical Analysis

Constructing the transdiagnostic internalizing factor

To create a comprehensive measure of internalizing psychopathology, we employed Bi-ESEM[77] using the WLSMV estimator to accommodate ordinal data (the detailed mathematical formulation of the model is provided in Supplementary Methods). Prior to modeling, we normalized all variables and conducted Kaiser-Meyer-Olkin (KMO) tests[78], retaining 21 items with values greater than 0.7. Model fit was assessed using the CFI, TLI, and the RMSEA, with benchmarks for good fit being CFI and TLI ≥ 0.95 and RMSEA ≤ 0.06 as per Hu and Bentler's guidelines[79]. The final model identified a general factor capturing shared variance across anxiety, depression, neuroticism, and sleep disturbances, along with specific factors representing unique symptom domains. To establish the robustness of the transdiagnostic internalizing factor model, we conducted analyses across four time points. Baseline measurements were used for predicting chronic diseases, and we validated the model structure using follow-up data collected at three additional time points: time point 1 (2011–2012), time point 2 (2014–2020), and time point 3 (2019+). Notably, the data collection during 2014–2020 coincided with the imaging assessments.

Prospective association between the transdiagnostic internalizing factor and chronic disease

Given the nature of the outcome variables and their suitability for survival analysis, we performed separate Cox proportional hazards regression models[80] to examine the prospective associations of transdiagnostic internalizing factor scores with 28 incident chronic disease outcomes, based on participants with complete data on internalizing symptoms, covariates, and health outcomes. The follow-up period extended from 2 years after baseline to April 1, 2022. Disease occurrence was coded as a binary variable (1 = diagnosed, 0 = not diagnosed). The observation time for each disease was defined as the number of days from two years after baseline to either the date of first diagnosis or the censoring date (April 1, 2022), whichever came first. We calculated hazard ratios (HRs) with 95% confidence intervals for 28 chronic physical diseases, applying FDR correction for multiple comparisons (significance threshold: $P < 0.05$). For 21 diseases demonstrating significant associations, we performed quantitative comparisons between the extreme transdiagnostic internalizing factor groups. Using a 30th percentile cutoff, age- and gender-matched participants were stratified into top and bottom groups. The mean age of first disease occurrence during the 14-year observation period was calculated for both groups, with statistical significance of intergroup differences assessed using independent samples t-tests. Additionally, the 21 diseases were categorized into six biological systems (neurological, musculoskeletal, cardiometabolic, respiratory, immune, and miscellaneous) for system-level analyses.

Cox regression models were used to estimate hazard ratios for each system, while cumulative incidence curves were employed to visualize group differences between the top and bottom percentiles.

Association between the transdiagnostic internalizing factor and multisystem biological indicators

To investigate potential biological mechanisms, we used general linear models to examine associations between transdiagnostic internalizing factor and biological indicators (233 brain grey matter volumes, 82 cardiovascular, 18 immune, 9 metabolic, 4 respiratory, and 17 musculoskeletal markers), calculating standardized regression coefficients (β_{std}). To investigate potential convergent biological mechanisms underlying the association between the transdiagnostic internalizing factor and chronic disease multimorbidity, we computed a multimorbidity measure based on the count of 21 chronic diseases that were significantly associated with the internalizing factor. The multimorbidity measure was then associated with the same 363 biomarkers using the analytical approach applied to transdiagnostic internalizing factors. In both sets of analyses, we adjusted for the same covariates and applied FDR correction to account for multiple comparisons.

Mediation analyses

We performed mediation analyses to assess how multisystem biomarkers mediate the longitudinal associations between the transdiagnostic internalizing factor and chronic disease incidence across six biological systems. To reduce model complexity, we extracted latent variables representing system-level biomarker profiles. For biological indicators significantly associated with both the transdiagnostic internalizing factor and the multimorbidity measure, we conducted system-specific principal component analysis (PCA), retaining the first principal component (PC1) to represent each system's biomarker profile. The proportion of variance explained by PC1 exceeded 45% across all systems, ranging from 46% for the musculoskeletal system to 85% for the respiratory system, with values of 75% for the brain, 72% for the metabolic system, 70% for the cardiovascular system, and 51% for the immune system. It should be noted that, because only a small number of participants had both common measures (e.g., grip strength, bone density) and abdominal MRI data, the musculoskeletal system PCA did not include abdominal MRI indicators. In our mediation models, the baseline transdiagnostic internalizing factor served as the independent variable, the PC1 of biological indicators as the mediator, and system-level chronic disease status (a binary variable) as the dependent variable. These models combined linear regression for transdiagnostic internalizing factor-to-biomarker paths with Cox regression for disease-related paths. We determined the significance of mediating effects based on 5,000 bootstrap iterations, applying FDR correction for multiple comparisons.

Data availability

This study was conducted using data from the UK Biobank (application number 85139). Access to the UK Biobank resource is available through a standard application process, subject to approval and associated access fees. Researchers can apply via the Access Management System (AMS) at https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access.

Declarations

Code availability

The analyses in the current manuscript were conducted using Python (version 3.10) and Mplus. Specifically, Python was utilized for statistical analyses—including Cox proportional hazards regression, generalized linear models, and mediation analyses—as well as for data visualization using the Matplotlib and Seaborn libraries. Mplus was employed for the Bi-ESEM to construct the transdiagnostic internalizing factor. Custom scripts and codes required to reproduce the main findings and figures are available from the corresponding author upon reasonable request.

Acknowledgements

This research was conducted using the UK Biobank Resource under Application Number 85139. We sincerely thank all the participants for generously contributing their time to this project and the UK Biobank team for their invaluable efforts in collecting, processing, and disseminating the data utilized in this study. This work is supported by STI2030-Major Projects (2022ZD0211900 to A.L.), the Natural Science Foundation of China (82171543 to A.L., 82372049 to B.L.), the Beijing Nova Program (20230484425 to B.L., 20250484761 to A.L) and the National Key R&D Program of China (2023YFC3603500 to Jiahong Wang).

Author contributions:

Conceptualization: A.L., Y.-N.H., T.G.

Methodology: Y.-N.H., T.G., A.L., B.L., X.-Y.L., K.-X.L.

Data Analysis & Preprocessing: Y.-N.H., T.G., J.-X.X., Y.-J.P., F.-Q.X., C.-S.D., S.-Z.H., X.-H.T., C.-Y.D.

Critical Feedback: J.-H.W., Y.Y., L.W., Q.L., Y.-H.H.

Writing: Y.-N.H., T.G., A.L., J.X.,

Supervision: B.L., J.X., A.L.

Competing interests: No competing interests.

References

1. Prince M et al (2007) No health without mental health. Lancet 370:859–877
2. Tian YE, Cole JH, Bullmore ET, Zalesky A (2024) Brain, lifestyle and environmental pathways linking physical and mental health. Nat Mental Health 2:1250–1261
3. Smagula SF, Stone KL, Fabio A, Cauley JA (2016) Risk factors for sleep disturbances in older adults: evidence from prospective studies. Sleep Med Rev 25:21–30

4. Jiang R et al (2024) The brain structure, inflammatory, and genetic mechanisms mediate the association between physical frailty and depression. Nat Commun 15:4411
5. Zhao Y et al (2023) The brain structure, immunometabolic and genetic mechanisms underlying the association between lifestyle and depression. Nat Mental Health 1:736–750
6. Kim H et al (2021) Internalizing psychopathology and all-cause mortality: a comparison of transdiagnostic vs. diagnosis-based risk prediction. World Psychiatry 20:276–282
7. Fleetwood KJ et al (2025) Depression and physical multimorbidity: A cohort study of physical health condition accrual in UK Biobank. PLoS Med 22:e1004532
8. De Hert M et al (2011) Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. World psychiatry 10:52
9. Correll CU, Detraux J, De Lepeleire J, De Hert M (2015) Effects of antipsychotics, antidepressants and mood stabilizers on risk for physical diseases in people with schizophrenia, depression and bipolar disorder. World psychiatry 14:119–136
10. Krueger RF, Eaton NR (2015) Transdiagnostic factors of mental disorders. World Psychiatry 14:27
11. Vaccarino V, Bremner JD (2024) Stress and cardiovascular disease: an update. Nat Reviews Cardiol 21:603–616
12. Won E, Kim Y-K (2016) Stress, the autonomic nervous system, and the immune-kynurenine pathway in the etiology of depression. Curr Neuropharmacol 14:665–673
13. Dimsdale JE (2008) Psychological stress and cardiovascular disease. J Am Coll Cardiol 51:1237–1246
14. Ziegler MG (2012) Psychological stress and the autonomic nervous system. In: Robertson D et al (eds) Primer on the Autonomic Nervous System. Academic, pp 291–293
15. Carney RM, Freedland KE, Veith RC (2005) Depression, the autonomic nervous system, and coronary heart disease. Psychosom Med 67:S29–S33
16. Janczura M et al (2020) Potential roles of psychological and oxidative stress in insulin resistance: a cohort-based study. Diabetol Metab Syndr 12:58
17. Majnarić LT et al (2021) Low psychological resilience in older individuals: an association with increased inflammation, oxidative stress and the presence of chronic medical conditions. Int J Mol Sci 22:8970
18. McCracken C et al (2022) Multi-organ imaging demonstrates the heart-brain-liver axis in UK Biobank participants. Nat Commun 13:7839
19. Williams LM (2016) Precision psychiatry: a neural circuit taxonomy for depression and anxiety. Lancet Psychiatry 3:472–480
20. Schmaal L et al (2017) Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. Mol Psychiatry 22:900–909

21. Watson D et al (2022) Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. World Psychiatry 21:26–54
22. Kotov R et al (2017) The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. J Abnorm Psychol 126:454
23. Lahey BB, Moore TM, Kaczkurkin AN, Zald DH (2021) Hierarchical models of psychopathology: Empirical support, implications, and remaining issues. World Psychiatry 20:57–63
24. Widiger TA et al (2019) Personality in a hierarchical model of psychopathology. Clin Psychol Sci 7:77–92
25. DeYoung CG et al (2022) The distinction between symptoms and traits in the Hierarchical Taxonomy of Psychopathology (HiTOP). J Pers 90:20–33
26. Snyder HR, Young JF, Hankin BL (2017) Strong homotypic continuity in common psychopathology-, internalizing-, and externalizing-specific factors over time in adolescents. Clin Psychol Sci 5:98–110
27. Eaton NR, Krueger RF, Oltmanns TF (2011) Aging and the structure and long-term stability of the internalizing spectrum of personality and psychopathology. Psychol Aging 26:987–993
28. Tateno A, Nogami T, Sakayori T, Yamamoto K, Okubo Y (2023) Depression as a prodromal symptom of neurodegenerative diseases. J Nippon Med school 90:157–164
29. Shdo SM et al (2020) Depressive symptom profiles predict specific neurodegenerative disease syndromes in early stages. Front Neurol 11:446
30. Freitas C, Deschênes S, Au B, Smith K, Schmitz N (2016) Evaluating lifestyle and health-related characteristics of older adults with co-occurring depressive symptoms and cardiometabolic abnormalities. Int J Geriatr Psychiatry 31:66–75
31. Gong L et al (2022) Association between single and multiple cardiometabolic diseases and depression: a cross-sectional study of 391,083 participants from the UK biobank. Front Public Health 10:904876
32. Leonard B (2000) Stress, depression and the activation of the immune system. World J Biol Psychiatry 1:17–25
33. Aich P, Potter AA, Griebel PJ (2009) Modern approaches to understanding stress and disease susceptibility: A review with special emphasis on respiratory disease. Int J Gen Med, 19–32
34. Bair MJ, Wu J, Damush TM, Sutherland JM, Kroenke K (2008) Association of depression and anxiety alone and in combination with chronic musculoskeletal pain in primary care patients. Psychosom Med 70:890–897
35. Berk M et al (2023) Comorbidity between major depressive disorder and physical diseases: a comprehensive review of epidemiology, mechanisms and management. World psychiatry 22:366–387
36. Kobrosly RW, Seplaki CL, Cory-Slechta DA, Moynihan J, van Wijngaarden E (2013) Multisystem physiological dysfunction is associated with depressive symptoms in a population-based sample of older adults. Int J Geriatr Psychiatry 28:718–727

37. Wilcox CE, Pommy JM, Adinoff B (2016) Neural circuitry of impaired emotion regulation in substance use disorders. Am J Psychiatry 173:344–361
38. Fettes P, Schulze L, Downar J (2017) Cortico-striatal-thalamic loop circuits of the orbitofrontal cortex: promising therapeutic targets in psychiatric illness. Front Syst Neurosci 11:25
39. Zhao G et al (2009) Depression and anxiety among US adults: associations with body mass index. Int J Obes 33:257–266
40. Zasadzka E, Pieczyńska A, Trzmiel T, Kleka P, Pawlaczyk M (2021) Correlation between handgrip strength and depression in older adults—a systematic review and a meta-analysis. Int J Environ Res Public Health 18:4823
41. Zhao B et al (2023) Heart-brain connections: Phenotypic and genetic insights from magnetic resonance images. Science 380:abn6598
42. Mello Ad. A. F. d., Mello MFd, Carpenter LL, Price LH (2003) Update on stress and depression: the role of the hypothalamic-pituitary-adrenal (HPA) axis. *Brazilian Journal of Psychiatry* 25, 231–238
43. Guilliams TG, Edwards L (2010) Chronic stress and the HPA axis. standard 9:1–12
44. Sharan P, Vellapandian C (2024) Hypothalamic-pituitary-adrenal (HPA) axis: Unveiling the potential mechanisms involved in stress-induced Alzheimer's disease and depression. Cureus 16
45. Rusch JA, Layden BT, Dugas LR (2023) Signalling cognition: the gut microbiota and hypothalamic-pituitary-adrenal axis. Front Endocrinol 14:1130689
46. Knezevic E, Nenic K, Milanovic V, Knezevic N (2023) N. The role of cortisol in chronic stress, neurodegenerative diseases, and psychological disorders. Cells 12:2726
47. Ali MK, Chen JD (2023) Roles of heart rate variability in assessing autonomic nervous system in functional gastrointestinal disorders: a systematic review. Diagnostics 13:293
48. Anaabawati MA, Rumahorbo H, Pujiastuti RSE (2021) Effectiveness of cognitive behavioral therapy and muscle relaxation on fasting blood sugar levels among type 2 diabetes mellitus: A systematic review. Int J Nurs Health Serv (IJNHS) 4:153–161
49. Borovac JA, D'Amario D, Bozic J, Glavas D (2020) Sympathetic nervous system activation and heart failure: Current state of evidence and the pathophysiology in the light of novel biomarkers. World J Cardiol 12:373
50. Gale CR, Martyn CN, Cooper C, Sayer AA (2007) Grip strength, body composition, and mortality. Int J Epidemiol 36:228–235
51. Mahmood A et al (2023) Neuroticism personality traits are linked to adverse cardiovascular phenotypes in the UK Biobank. Eur Heart Journal-Cardiovascular Imaging 24:1460–1467
52. Piché M-E, Tchernof A, Després J-P (2020) Obesity phenotypes, diabetes, and cardiovascular diseases. Circul Res 126:1477–1500
53. Batty GD, Gale CR, Kivimäki M, Deary IJ, Bell S (2020) Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. BMJ 368:m441

54. Kim SM, Hong JS, Min KJ, Han DH (2019) Brain functional connectivity in patients with somatic symptom disorder. Biopsychosoc Sci Med 81:313–318
55. Moloney RD, Desbonnet L, Clarke G, Dinan TG, Cryan JF (2014) The microbiome: stress, health and disease. Mamm Genome 25:49–74
56. Palmer LJ (2007) UK Biobank: bank on it. Lancet 369:1980–1982
57. Mutz J, Roscoe CJ, Lewis CM (2021) Exploring health in the UK Biobank: associations with sociodemographic characteristics, psychosocial factors, lifestyle and environmental exposures. BMC Med 19:240
58. Littlejohns TJ et al (2020) The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. Nat Commun 11:2624
59. Hussain S et al (2020) Depressive and anxiety symptomatology among people with asthma or atopic dermatitis: A population-based investigation using the UK Biobank data. Brain Behav Immun 90:138–144
60. Eysenck SB, Eysenck HJ, Barrett P (1985) A revised version of the psychoticism scale. Pers Indiv Differ 6:21–29
61. Nagel M, Watanabe K, Stringer S, Posthuma D, Van Der Sluis (2018) Item-level analyses reveal genetic heterogeneity in neuroticism. Nat Commun 9:905
62. Sambou ML et al (2021) Associations between sleep quality and health span: a prospective cohort study based on 328,850 UK biobank participants. Front Genet 12:663449
63. Zhang Y et al (2023) Identifying modifiable factors and their joint effect on dementia risk in the UK Biobank. Nat Hum Behav 7:1185–1195
64. Chudasama YV et al (2020) Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: a longitudinal cohort study. PLoS Med 17:e1003332
65. GBD 2019 Mental Disorders Collaborators (2022) Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 9:137–150
66. Gold SM et al (2020) Comorbid depression in medical diseases. Nat reviews Disease primers 6:69
67. Gaser C et al (2024) CAT: a computational anatomy toolbox for the analysis of structural MRI data. Gigascience 13:giae049
68. Ashburner J (2007) A fast diffeomorphic image registration algorithm. NeuroImage 38:95–113
69. Fan L et al (2016) The human brainnetome atlas: a new brain atlas based on connectional architecture. Cereb Cortex 26:3508–3526
70. Diedrichsen J, Balsters JH, Flavell J, Cussans E (2009) Ramnani, N. A probabilistic MR atlas of the human cerebellum. NeuroImage 46:39–46
71. Neudorfer C et al (2020) A high-resolution in vivo magnetic resonance imaging atlas of the human hypothalamic region. Sci data 7:305

72. Petersen SE et al (2016) UK Biobank's cardiovascular magnetic resonance protocol. J Cardiovasc Magn Reson 18:8
73. Bai W et al (2018) Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. J Cardiovasc Magn Reson 20:65
74. Zhang XY et al (2014) Full blood count parameters for the detection of asthma inflammatory phenotypes. Clin experimental allergy 44:1137–1145
75. Fest J et al (2018) Reference values for white blood-cell-based inflammatory markers in the Rotterdam Study: a population-based prospective cohort study. Sci Rep 8:10566
76. Milaneschi Y et al (2021) Association of inflammation with depression and anxiety: evidence for symptom-specificity and potential causality from UK Biobank and NESDA cohorts. Mol Psychiatry 26:7393–7402
77. Jennrich RI, Bentler PM (2012) Exploratory bi-factor analysis: The oblique case. Psychometrika 77:442–454
78. Williams B, Onsman A, Brown T (2010) Exploratory factor analysis: A five-step guide for novices. Australasian J paramedicine 8:1–13
79. Hu LT, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct equation modeling: multidisciplinary J 6:1–55
80. Andersen PK, Gill RD (1982) Cox's regression model for counting processes: a large sample study. annals Stat 10:1100–1120

Figures

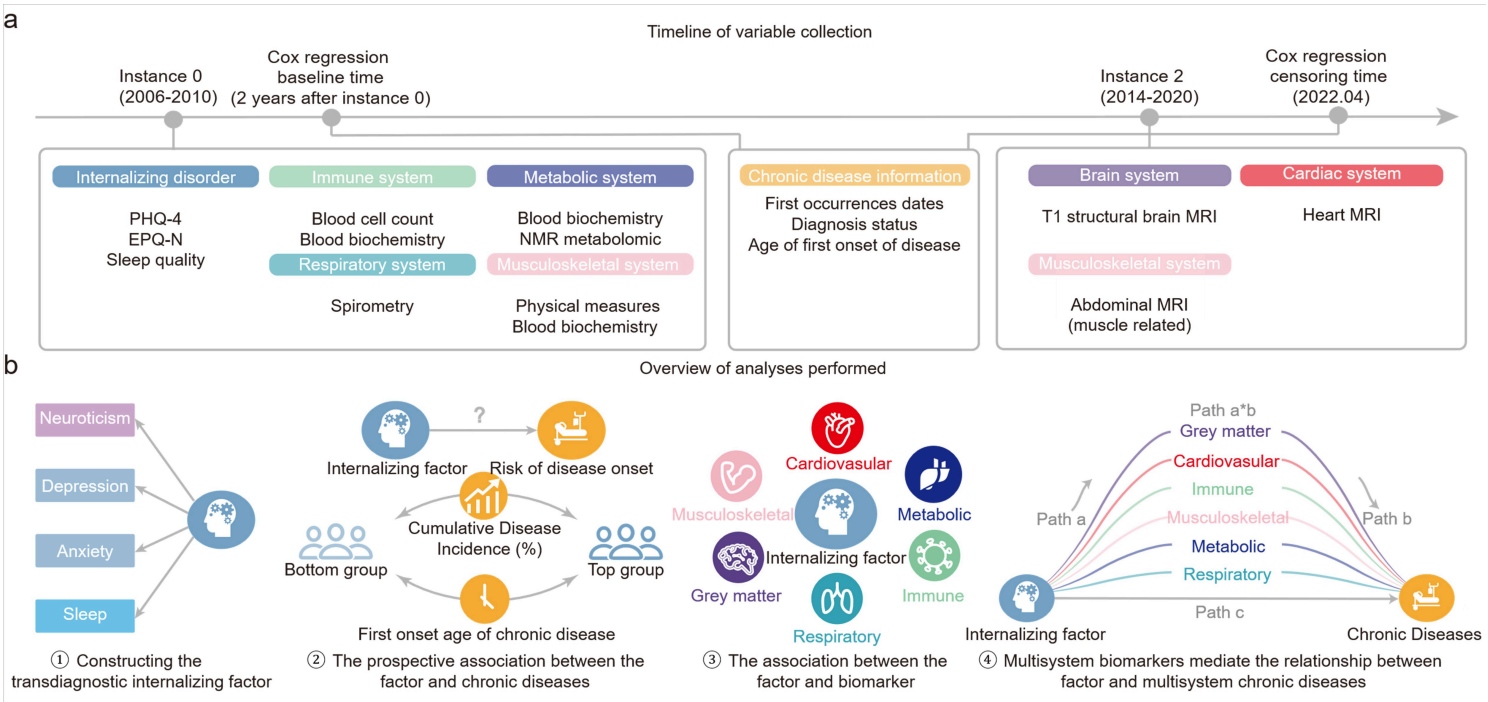

Figure 1

Study design and analytical framework. (a) Timeline of data collection and follow-up. Instance 0 (2006-2010): internalizing disorder symptoms (PHQ-4, EPQ-N, sleep quality), immune system biomarkers (blood cell count, biochemistry), metabolic system biomarkers (blood biochemistry, NMR metabolomics), respiratory system biomarkers (spirometry), and musculoskeletal system biomarkers (physical measures, blood biochemistry); Instance 2 (2014-2020): brain imaging (T1 MRI), cardiac imaging (heart MRI), and musculoskeletal imaging (abdominal MRI). Cox regression baseline (2 years after Instance 0): defined as the reference point for assessing disease risk. Cox regression censoring (April 1, 2022): endpoint for tracking disease incidence. Between the Cox regression baseline and cox regression censoring, disease status, age of first onset, and first incidence of diseases were assessed. **(b)** Overview of analytical framework. ① Construction of the transdiagnostic internalizing factor from four dimensions (depression, anxiety, neuroticism, and sleep disturbances).②Prospective associations with incident chronic diseases: using Cox regression models (baseline: 2 years post-Instance 0; censoring: April 2022), we examined how the transdiagnostic internalizing factor predicts first occurrence of 28 chronic diseases, including hazard ratios for disease incidence, differences in age at first diagnosis, and cumulative disease burden across six physiological systems.③ Associations with multi-system biomarkers across neurological and five peripheral systems. ④ Mediation analyses: testing whether multi-system biomarkers mediate the relationship between the internalizing factor and chronic disease risk. Abbreviation: PHQ-4, The Patient Health Questionnaire-4; EPQ-N, The Eysenck Personality Questionnaire-Revised Short Form.

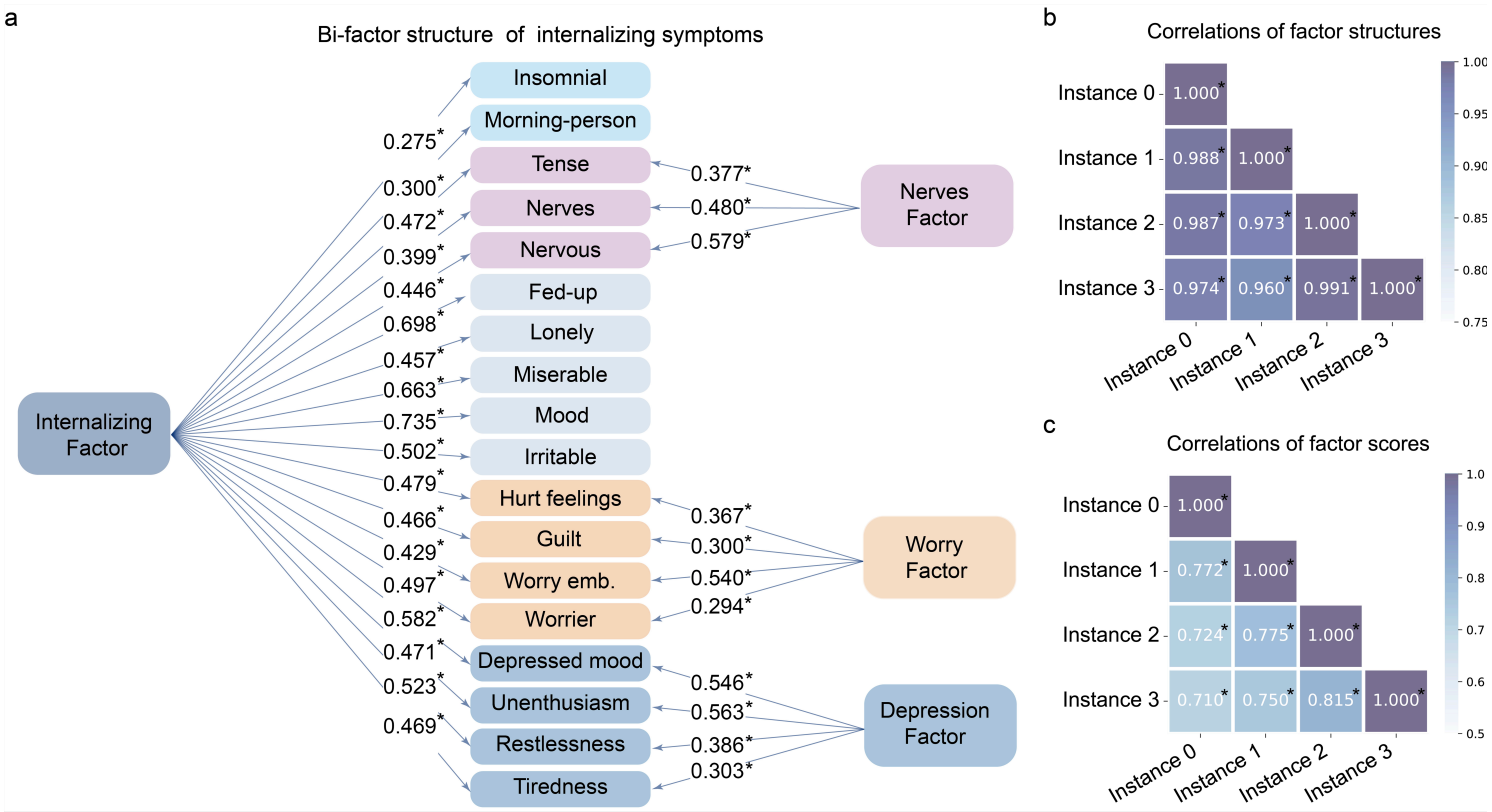

Figure 2

Bi-factor structure and temporal stability of the transdiagnostic internalizing factor. (a) Baseline factor structure (n = 158,435): Bi-ESEM identified one general transdiagnostic internalizing factor and three specific factors (nerves, worry, depression) across 18 symptom items in 158,435 participants. Only factor loadings >0.25 are displayed. **(b)** Structural stability: Factor loading patterns remained highly consistent across four waves over 16 years (Pearson r between loading matrices: 0.96–0.99), indicating measurement invariance. **(c)** Individual score stability: Participants' factor scores showed strong test-retest correlations (r = 0.71–0.82), reflecting stable rank-ordering of individuals over time. * indicates false discovery rate (FDR) corrected$P$ < 0.05.

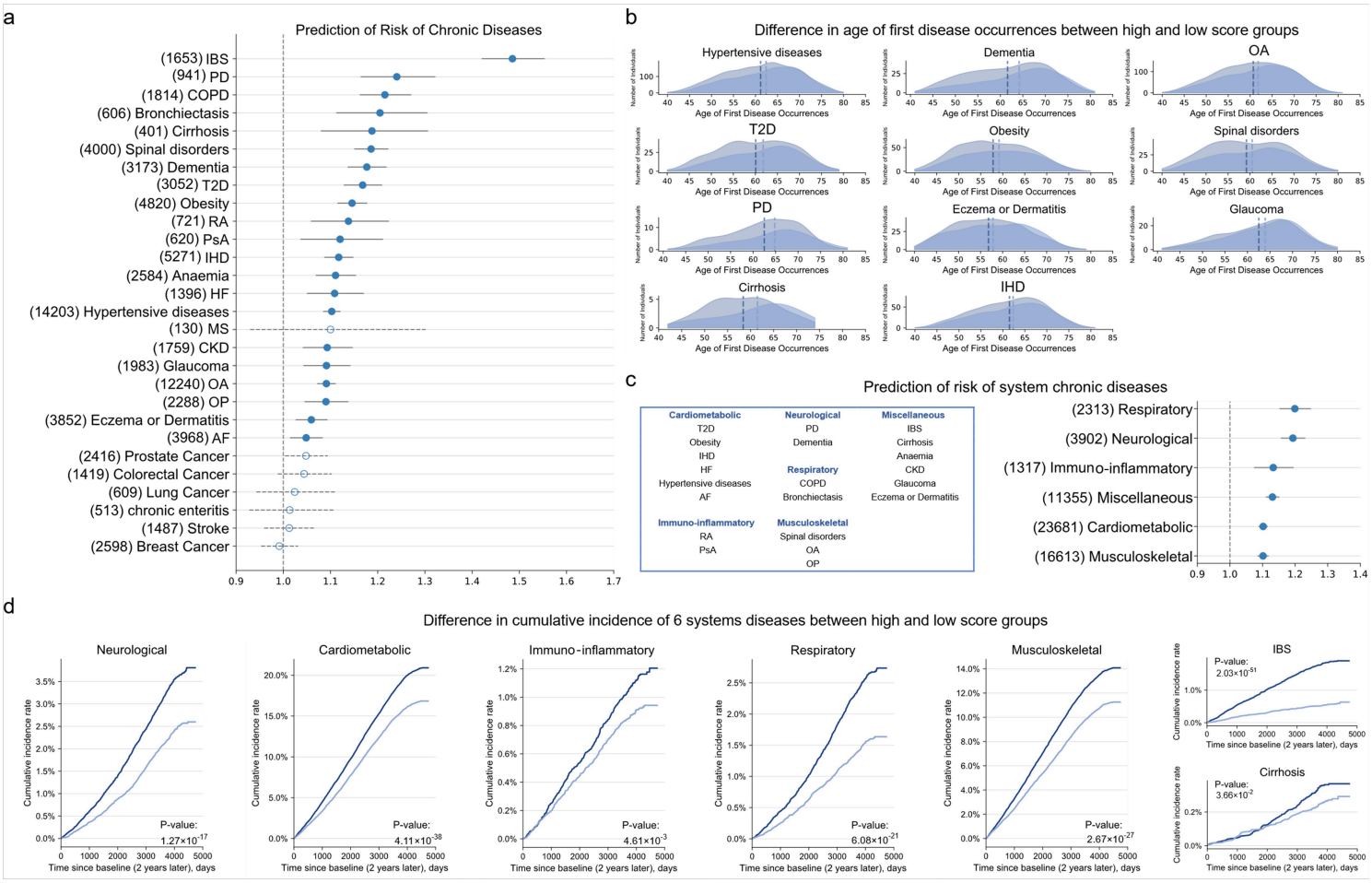

Figure 3

Prospective associations between the transdiagnostic internalizing factor and chronic diseases. (a) Cox regression analyses showing hazard ratios (HRs) for 28 chronic diseases; the number of cases is shown in parentheses to the left of each disease. Solid circles and lines indicate FDR-significant associations, while hollow circles and dashed lines indicate non-significant results. **(b)**Differences in age at first onset between high- and low-score groups for 21 significant diseases; high-score individuals developed cirrhosis, dementia, and Parkinson's disease 2–3 years earlier on average. **(c)** System-level analyses across six categories revealed the strongest associations for respiratory and neurological diseases. Case numbers are shown in parentheses. **(d)**Cumulative incidence curves show consistently higher disease rates in the high-score group across all systems during 14 years of follow-up.

Abbreviation: IBS, irritable bowel syndrome; COPD, chronic obstructive pulmonary disease; T2D, type 2 diabetes; RA, rheumatoid arthritis; PsA, psoriatic arthritis; IHD, ischemic heart disease; HF, heart failure; MS, multiple sclerosis; CKD, chronic kidney disease; OA, osteoarthritis; OP, osteoporosis; AF, atrial fibrillation.

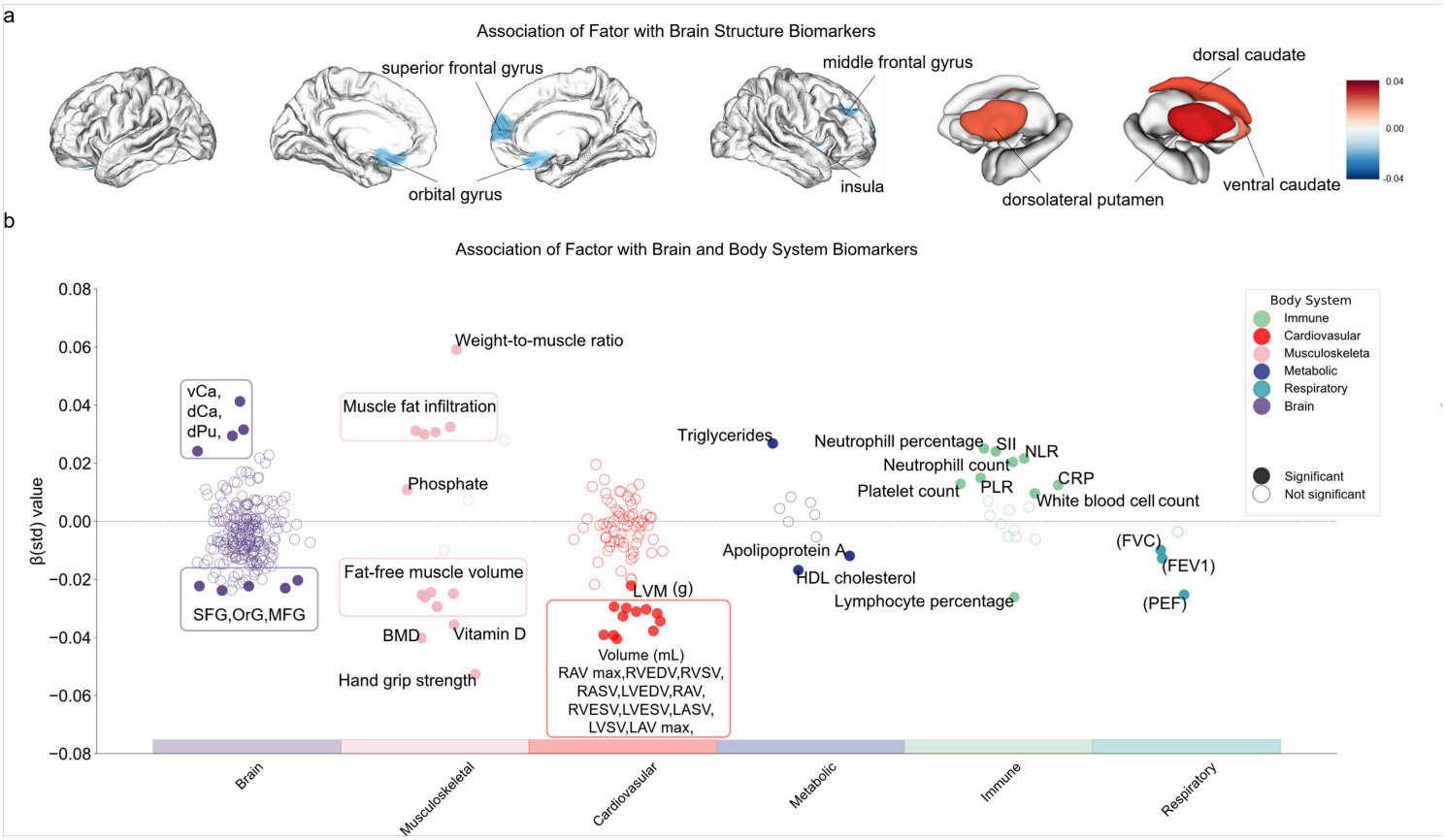

Figure 4

Associations between the transdiagnostic internalizing factor and biological markers across brain and body systems. (a) Brain regions whose grey matter volumes were significantly associated with the internalizing factor (n = 9,089), identified using general linear regression controlling for nine covariates. Red indicates positive associations and blue indicates negative associations. **(b)** Associations between the transdiagnostic internalizing factor and systemic biomarkers spanning the immune, cardiovascular, musculoskeletal, metabolic, and respiratory systems. The y-axis denotes standardized regression coefficients (*β*), and the x-axis lists individual biomarkers within each system. Solid circles indicate associations that remained significant after FDR correction, while open circles denote non-significant results. Abbreviations: SFG, superior frontal gyrus; OrG, orbital gyrus; MFG, middle frontal gyrus; vCa, ventral caudate; dCa, dorsal caudate; dPu, dorsal putamen; LVM, left ventricular mass; RAV max, right atrial maximum volume; RVEDV, right ventricular end-diastolic volume; RVSV, right ventricular stroke volume; RASV, right atrial stroke volume; LVEDV, left ventricular end-diastolic volume; RAV, right atrial volume; RVESV, right ventricular end-systolic volume; LVESV, left ventricular end-systolic volume; LASV, left atrial stroke volume; LVSV, left ventricular stroke volume; LAV max, left atrial maximum volume; CRP, c-reactive protein; PLR, platelet-to-lymphocyte ratio; NLR, neutrophil-to-lymphocyte ratio; SII, systemic

immune-inflammation index; BMD, bone mineral density; FVC, forced vital capacity; FEV1, forced expiratory volume in 1 second; PEF, peak expiratory flow.

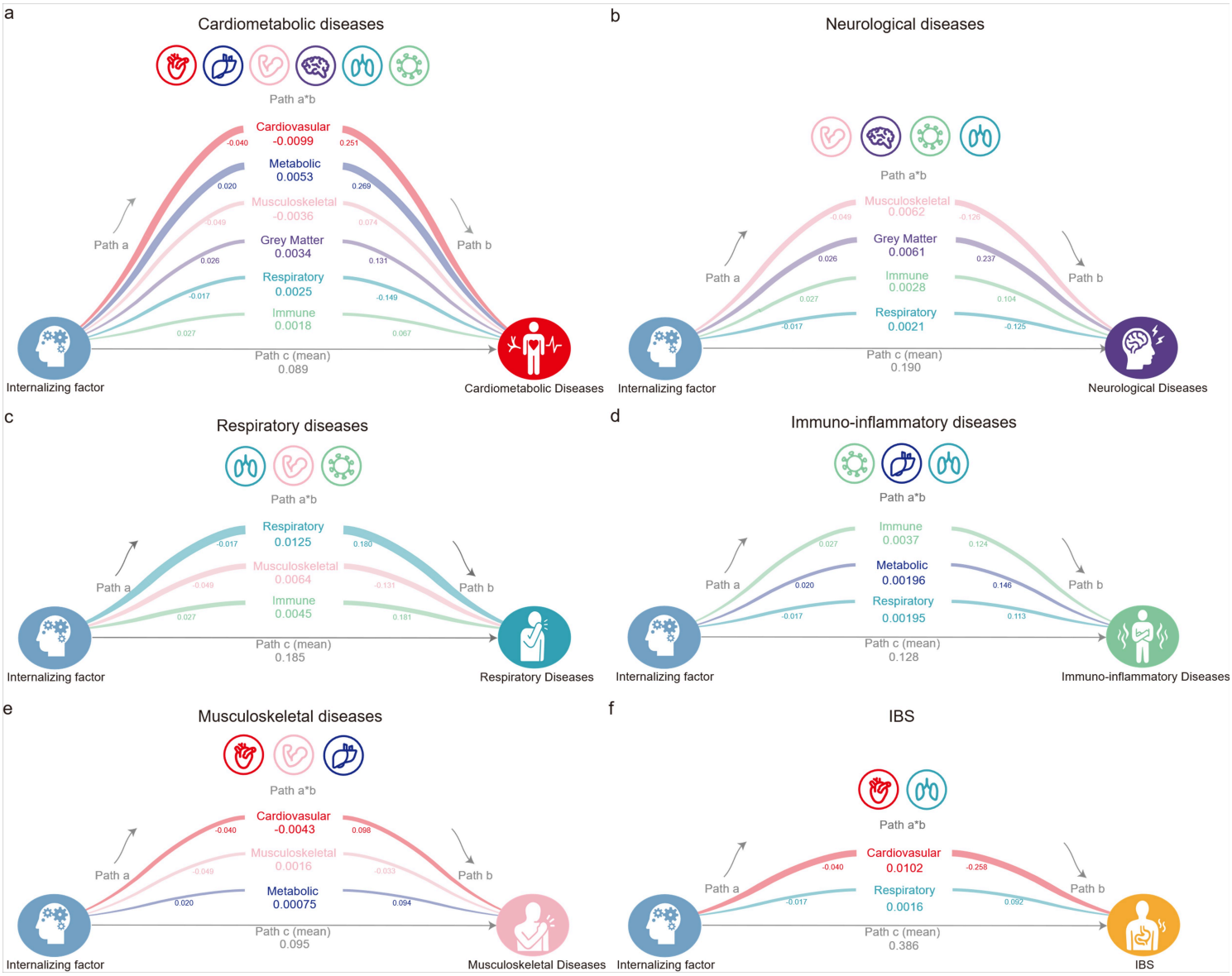

Figure 5

Mediation of the association between the transdiagnostic internalizing factor and chronic diseases by systemic biomarkers. Panels (a–f) illustrate the mediating roles of composite biomarker indices from multiple biological systems in linking the transdiagnostic internalizing factor (left) to chronic disease outcomes (right). Curved arrows indicate indirect paths, with colors denoting the contributing systems (grey matter, cardiovascular, immune, musculoskeletal, metabolic, respiratory; legends above each panel). In the mediation model, coefficients for path *a*are displayed on the left and for path *b* on the right. Standardized indirect effects (*a* × *b*) are indicated along the intermediate paths. The figure displays only mediation pathways that reached FDR significance. **(a)** Cardiometabolic diseases: mediated primarily through cardiovascular and metabolic markers. **(b)** Neurological diseases: mediated primarily via musculoskeletal and neurological pathways. **(c)**Respiratory diseases: mediated through respiratory,

musculoskeletal, and immune pathways. **(d)** Immuno-inflammatory diseases: mediated via immune, respiratory, and metabolic markers. **(e)** Musculoskeletal diseases: mediated mainly through cardiovascular and musculoskeletal markers. **(f)**IBS: mediated via cardiovascular and respiratory markers.

Supplementary Files

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

- TifactorsupplementaryNC.docx