Research Square

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

Molecular Heterogeneity of Aging: Insights from Integrated Plasma Proteome

Elizaveta Sarygina

lizalesa@gmail.com

Institute of Biomedical Chemistry

Varvara Blinova

Institute of Biomedical Chemistry

Arina Gordeeva

Institute of Biomedical Chemistry

Anna Kliuchnikova

Institute of Biomedical Chemistry

Anna Kozlova

Institute of Biomedical Chemistry

Svetlana Tarbeeva

Institute of Biomedical Chemistry

Aziz Abdurahimov

Institute of Biomedical Chemistry

Ekaterina Ilgisonis

Institute of Biomedical Chemistry

Svetlana Novikova

Institute of Biomedical Chemistry

Viktor Zgoda

Institute of Biomedical Chemistry

Andrey Lisitsa

Institute of Biomedical Chemistry

Alexander Gusev

Institute of Biomedical Chemistry

Irina Strazhesko

National Medical Research Center for Preventive Medicine (NMRCPM)

Alexey Churov

Pirogov Russian National Research Medical University

Olga Tkacheva

Pirogov Russian National Research Medical University

Vadim Govorun

Scientific Research Institute of Systems Biology and Medicine

Alexander Archakov

Institute of Biomedical Chemistry

Elena Ponomarenko

Institute of Biomedical Chemistry

Research Article

Keywords: biological ageing, plasma proteome, interindividual variability, biomarkers, Selected Reaction Monitoring, cystatin C, complement system, endocrine regulation, machine learning

Posted Date: August 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10578768/v1

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

Additional Declarations: No competing interests reported.

Abstract

We conducted an integrative analysis of the plasma proteome and routine clinical blood parameters to identify biomarkers of physiological aging in a cross-sectional cohort of 99 healthy individuals (aged 18–90 years) systematically recruited within the national "RUSS-AGE" project. Plasma proteomic profiling was performed using targeted Selected Reaction Monitoring mass spectrometry with 49 stable isotope-labelled standards, alongside comprehensive clinical biochemistry tests.

Our population-level analysis identified strong age-associated correlations for Somatomedin C (IGF-1), Dehydroepiandrosterone sulphate (DHEA-S), and Cystatin C (CST3). Furthermore, we defined an eight-marker signature that reliably distinguishes between younger and older individuals using ML models. This signature indicates a systemic shift from a metabolically active, protective state to one characterized by chronic inflammation and reduced resilience, aligning with established hallmarks of aging.

Critically, we found that the inter-individual variability of measured parameters is age-dependent. Protein levels exhibited peak variability in the 28–45 age group, while most clinical parameters showed maximal dispersion later in life (61–80 years). This high variability underscores the following: some older individuals maintain a molecular profile typical of younger population norms, and vice versa. Collectively, integrative plasma proteomic and clinical blood analysis reveal molecular heterogeneity of human aging. These findings may inform future strategies for monitoring and promoting healthy aging on an individualized basis.

Introduction

The identification of robust biomarkers that capture the heterogeneity of human biological aging is crucial for developing personalized healthspan interventions. Given that a lot of researchers are concentrated on this problem, numerous definitions of the term “aging” have arisen. In most cases, aging is understood as a natural, age-dependent process of irreversible deterioration of the physiological functions of the organism. It is characterized by the accumulation of various disturbances at the molecular and cellular levels, which are ultimately associated with an increased risk of disease and death[1].

It is generally accepted that increase of chronological age may be represented by the "older age phenotype" determined by a profile of biological parameter[2]. A large number of tests for determination of biological age have been developed. Various types of “aging clocks” have been described - transcriptomic, epigenetic, proteomic, associated with the microbiome, inflammation, shifts in the immune system, neuroimaging, etc. [3–5]. The main idea that underlies the development of such clocks is the simplification of biological age calculation. Although certain patterns can be indeed found, some parameters violate the established correlations and represent outliers. Moreover, in real life these correlations are often disrupted: some individuals at the age of 60 may operate better (and therefore live

for several more decades) than the average 40-year-old individual, who is likely to die within the two-year period. Therefore, results from the previous findings indicate the limited applicability of aging clocks[6].

The underlying cause of this limited applicability becomes apparent when considering the aging process from the individual-level perspective. The relevance of biological age determination from a single time-point analysis is questionable. Numerous factors, such as stress exposure, lifestyle changes, and circadian rhythm disruptions, may cause short-term fluctuations in certain parameters, which would affect the interpretation of the determined biological age results[7–9].

Furthermore, the feasibility of extrapolating the data underlying aging clocks to an individual organism also raises concerns. The majority of cohorts used in developed clocks represent European population[4]. This results in underrepresentation of other ethnic groups and leads to the above mentioned restricted applicability. Moreover, proposed clocks do not capture critical organism-specific dynamics: the individual’s homeostatic variability, regulatory flexibility, and compensatory capacity. Consequently, the vast majority of existing population-specific tests provide an assessment of the organism's state across multiple levels at a given time but do not consider adaptability, resilience, and the margin of safety of individual human resources.

However, the reserve capacity of an individual may be considered as one of the main intrinsic factors regulating the aging process. Therefore, the key characteristic that may distinguish an elderly organism from a young one is not the chronological age, but the increased vulnerability, or fragility, of the system[10], i.e., a decrease in the ability to adapt to different types of loads and to regain the homeostasis. In this regard, aging could be defined as a time-dependent function describing the mobilization of physiological reserves in counteracting destabilizing factors. This function may serve as an indicator of viability and as a predictor of one's resources exhaustion. Framework of adaptive potential and compensatory capacity requires determination and further elucidation, as it has a potential to define the predictive model of individual resilience.

Quantifying this internal resource and revealing what underlies this term on a molecular basis presents a major challenge. For this purpose, the proteome emerges as a prime candidate, given its capacity to integrate multi-system signals and represent the organism's functional state[11]. Plasma proteins are exceptionally informative, acting as the final functional effectors in the body. Modern analytical technologies, such as mass spectrometry (MS) and aptamer- or antibody-based platforms, allow for the simultaneous quantification of numerous circulating proteins[12]. Large studies have revealed complex, nonlinear proteome changes with age, including "waves" of remodeling during specific decades of life and marked sex differences in aging trajectories[13–15]. Biological significance of proteins, combined with the practical advantages of accessible sampling and robust analytical techniques, makes these molecules a good basis for identifying age-associated changes of homeostasis in various age groups[16].

This study identified age-associated molecular parameters that form an integral index of frailty. Proteomic measurements were used in combination with clinical and laboratory parameters. We

conducted a targeted MS analysis of plasma proteins in a cohort of apparently healthy donors spanning a broad age range. To achieve this, we employed a targeted panel with stable isotope-labelled standards in combination with advanced statistical selection methods and predictive modelling.

Results

Age-related dynamics and sex-specific features of blood parameters

Each person possesses a chronological age and a unique life history, indirectly mirrored in a set of clinical, laboratory, molecular, physical, and cognitive parameters. The internal resources of the body differ among individuals and are not amenable to the measurement. Their assessment is possible only indirectly, either via analysis of functional capabilities or, more objectively, via molecular markers. Molecular patterns can thus form the basis for quantifying the body's reserves.

We performed an integrated analysis of the plasma proteome and clinical blood parameters in a cross-sectional cohort of 99 healthy donors (aged 18–90), of whom 49 were women and 50 were men. A substantial proportion of participants were either young (below 30 years; 45%) or elderly (above 60 years; 35%), whereas the middle-aged group (30–50 years) was less represented (Fig. 1a). Plasma proteins were quantified using targeted Selected Reaction Monitoring mass spectrometry (SRM) with 49 stable isotope-labelled standards, alongside comprehensive clinical biochemistry. One protein was excluded from subsequent analysis because it was detected in only five individuals, resulting in an inadequate sample size for robust statistical assessment. Detailed demographic and clinical characteristics of cohort (n = 99), along with the complete list of 49 plasma proteins are provided in Table S1 - Supplementary.

Participant age ranged from 18 to 90 years, a mean was of 46.2 ± 18.7 years and a median of 44 years. Sex distribution within each age category was balanced and showed no statistically significant differences. In the female subgroup, the mean age was 45.8 ± 17.9 years (median 43 years), while in the male subgroup it was 46.6 ± 19.5 years (median 45 years) (Fig. 1b). The maximum age among women was 81 years, whereas in men it reached 90 years.

Notably, in 76% of participants the biological age predicted by the PhenoAge algorithm (Fig. 1c) was lower than their chronological age, consistent with previous observations in healthy populations[17]. Only 10% of participants showed higher biological than chronological age, which may reflect individual variation in ageing rate or unmeasured lifestyle and environmental factors (Fig. 1c). All participants underwent extended analysis of clinical and biochemical parameters, including major blood and serum indices (Table S1 - Supplementary). Baseline characteristics and comparative statistics across WHO age categories are summarized in Table S2 - Supplementary.

The cohort demonstrated general homogeneity across age groups for key hematological parameters, providing a reliable basis for analysis. As expected, older participants showed higher glucose, BMI, and alkaline phosphatase, alongside lower erythrocyte counts. Correlation analysis revealed a strong negative association between age and Somatomedin C ($r = -0.76$, $p < 0.05$), a ageing marker[18] (Fig. 1d).

Other markers like Cystatin C and urea increased with age, indicating changes in renal function[19,20] and metabolism. In contrast, DHEA-S levels declined sharply (see Fig. 2a): in older individuals they amounted to only 10–20% of those observed in youth[21]. This pattern characterized by reduced anabolic signals (Somatomedin C, DHEA-S) with increased catabolic markers (Cystatin C, UREA) is a basis for assessing biological age, see Fig. 2a-d.

Sex-stratified analysis revealed differences in the age-related sensitivity of laboratory parameters. In men the strongest associations with age were observed (see Fig. 2e) for insulin-like growth factor 1 (Somatomedin C; $r = -0.842$, $p = 1.72 \times 10^{-14}$), dehydroepiandrosterone sulphate (DHEA-S; $r = -0.767$, $p = 8.11 \times 10^{-11}$), and cystatin C ($r = 0.684$, $p = 4.29 \times 10^{-8}$). In women significant associations were mainly observed for metabolites of phenolic and organic acids, including p-hydroxyphenylpyruvate ($r = 0.443$, $p = 1.43 \times 10^{-3}$), urinary creatinine, 2-ketoglutaric acid, succinic acid, oxalic acid, glycolic acid, homogentisic acid, and 3-hydroxyisovaleric acid, reflecting sex-specific metabolic alterations associated with ageing.

Overall, although hormonal, metabolic, and immune parameters demonstrated age-associated changes in both sexes, the magnitude of these alterations was greater in men. Somatomedin C, DHEA-S, Cystatin C, UREA, and HbA1c showed the most significant between-group differences and can thus be considered as robust biochemical indicators of aging, with significant sex-specific variation.

A linear regression model was constructed to identify key predictors of chronological age. The high coefficient of determination ($R^2 \approx 0.9$) of the resulting model indicates that the selected set of biomarkers has high predictive power.

Based on the constructed regression model, the parameters with the greatest impact on age prediction were selected. A pattern of such predictors was identified, including progesterone, C-reactive protein levels, sex hormone-binding globulin (SHBG), and a number of organic acids (tricarballylic, acetoacetic, isocitric, and malic acids) (see Fig. S1a).

A separate analysis and ranking of the parameters by the degree of spread of values (variability, see Fig. S1b) was conducted. This allowed us to identify parameters with the greatest individual variability, which may be most informative for identifying aging heterogeneity. The list of the most variable parameters does not coincide with the top predictors based on the regression model results (see Fig. 1).

Vulnerability points may vary among individuals, and these factors shape individual aging trajectories. High variability in T-lymphocytes and aromatic acid metabolites suggests that the immune system, microbiome-dependent metabolism, and hematopoiesis may age at different rates in different

individuals. These variable signatures are promising candidates for identifying aging subtypes (endotypes) and developing personalized approaches.

Phenotypic convergence

Analysis of age-related trajectories for clinical laboratory and proteomic measures revealed a key phenomenon highlighting the heterogeneity of aging. Despite clear population-wide trends, we identified individuals from extreme age groups (young and old) who exhibited nearly identical molecular profiles. This pattern was observed in both proteomic data and routine clinical laboratory parameters. Figure 4 presents age-trajectories of clinical laboratory parameters (a, b) and plasma proteins measured by targeted proteomics (SRM) (c, d).

Somatomedin C and Cystatin C are represented as examples of clinical laboratory parameters that decrease (Somatomedin C) or increase with age (Cystatin C). The age-related trajectory of a protein that increases with age is illustrated using Cystatin C as an example. No protein among those analyzed demonstrated a decreasing concentration trend with age. For clarity, data points corresponding to samples from young (18–30 years) and old (70 + years) donors with statistically indistinguishable SRM-measured protein levels and concentrations of clinical laboratory parameters are circled in red and blue.

The data demonstrated a significant overlap between age groups for specific parameters, with some older subjects showing levels completely within the younger cohort's distribution, and conversely, some younger subjects displaying values characteristic of the older population.

These findings demonstrate that a distinct segment of the elderly population retains molecular characteristics statistically typical for much younger individuals. Therefore, our analysis of the employed proteomic and clinical laboratory panels underscores pronounced inter-individual heterogeneity in aging. The preservation of youthful profiles in late life, along with variable rates of decline across different systems, reveals that chronological age is a poor proxy for biological age. This principle is crucial for informing preventative strategies focused on enhancing resilience and prolonging healthspan.

Heterogeneity within age groups

Individual parameters or their combinations (e.g., levels of specific proteins, metabolites, or clinical measures) frequently exhibit substantial variability within a single age cohort (Fig. 5).

To precisely characterize the observed heterogeneity, we employed two complementary measures of variability: the coefficient of variation (CV) and the variance (see Methods). The CV offers a smoothed, summary view of relative variability across a group. The variance, however, captures the aggregated individual deviations, making it a more powerful metric for identifying age periods characterized by a

divergence in individual aging paths, as it is sensitive to the emergence of distinct subgroups (e.g., "fast" vs. "slow" agers) within a population.

Figure 5 represents the analysis employed variability metrics calculated from age-stratified cohorts balanced for sample size. In protein analysis, the CV shows a general declining trend with a minor elevation in the 28–45 age group (Fig. 5a). The variance of the same data also shows a peak at 28–45 years, while demonstrating a general declining trend reaching its minimum within 61–80 years (Fig. 5b). We observe that inter-individual variability for proteomic parameters changes across the lifespan. Notably, the decline of variability of proteins with age may indicate a reduction in the diversity of aging proteomic profiles.

When analyzing clinical laboratory parameters, the CV demonstrated no tendency to decline or increase (Fig. 5c). In contrast, the variance of the same data, while exhibiting an overall increase, displays a distinct peak within the 61–80 year range (Fig. 5d). This finding is consistent with data from other scientific groups reporting an age-related increase in the dispersion of molecular parameters[25,26].

We assume that our analysis may point to a previously unreported form of molecular divergence: a non-coincident peak in clinical laboratory variance alongside a trough in proteomic variability within the 61–80 year age group. We interpret such novel pattern as an evidence of a critical transitional phase and suggest that it might beanalogous to a developmental "crisis" point. In this period, clinical parameters reach a state of maximum molecular heterogeneity, while proteomic profiles conversely exhibit their greatest homogeneity. This fundamental dissociation suggests a significant, molecule-type-specific shift in biological trajectories during the transition into late adulthood, as evidenced even in a relatively non-large cohort .

A distribution of proteomic (Fig. 5e) and clinical laboratory (Fig. 5f) parameters by variability level were conducted. Using a color scheme (dark blue = very low, blue = low, yellow = moderate, red = high), we observe that Fig. 5e and Fig. 5f demonstrated a prevalence of the red color, indicating a dominance of molecules with high levels of variability. These results are in line with those shown in Fig. 5a,c. It is important to emphasize that the focus of this analysis is not on absolute quantities, but on the dynamics of inter-individual variability within age cohorts. On the protein level, we observed a prevalence of molecules with high levels of variability during youth (18–22 years) and peak within 28–45 years, with a declining trend afterwards. In the context of clinical laboratory parameters, no such trends or peaks were observed. This observation may indicate that the diversity of clinical laboratory molecular profiles does not systematically increase or decrease with age.

It should be mentioned that the divergence in variability between panels 5e and 5f may arise from platform scale: clinical assays leverage large epidemiological cohorts, while proteomic studies operate with smaller sample sizes. Platform-related batch effects or increased technical noise at larger scale may confound the assessment of biological heterogeneity. This observation is corroborated by principal component analysis, which reveals no clear clustering by chronological age, thereby demonstrating profound heterogeneity across the studied population (Fig. 6e,f).

Cluster analysis of clinical laboratory and proteomic data produced dendrograms that visually delineate age-related stratification within the study population. The clustering patterns indicate that both types of parameters capture underlying biological trajectories associated with aging (Fig. 6). Figure 6a presents the results of hierarchical clustering of samples based on clinical-laboratory parameters, while Fig. 6b shows clustering based on proteomic profiles. The leaves of the dendrogram are labeled according to donor age, with a color code reflecting age groups: from green (youngest) to red (long-lived individuals).

Clustering of clinical-laboratory data (Fig. 6a) revealed five main clusters, with long-lived individuals grouped into clusters #1 and #5. Comparison with proteomic data (Fig. 6b) showed only one discrepancy: two samples (from donors aged 26 and 45) were separated into a distinct outlier cluster (#3). The composition of clusters based on clinical laboratory indicators was heterogeneous with respect to age. For example, cluster 2 included both young donors (n = 6) and individuals older than 45 (n = 5). Cluster #3 contained eight long-lived individuals along with seven persons younger than 27. Cluster #4 consisted primarily of middle-aged individuals (28–45 years) but also included both long-lived donors and young donors (younger than 22 years).

Aforementioned findings indicate that the variability of the 182 clinical laboratory parameters does not allow for a clear separation of the sample into discrete age-based groups. This conclusion is supported by principal component analysis (PCA, Fig. 6e), which did not reveal discrete clusters: all samples, regardless of age group, were grouped together in the lower left quadrant of the plane.

Understanding the molecular context of aging requires a comparison of the clustering results from the two data types (Fig. 6a and 6b). The distance scales on both dendrograms are normalized to a range of 0 to 1. Differences in sample distribution between the two approaches are visualized using proportional percentage representations of group contributions to the clusters in Fig. 6c (clinical laboratory data) and 6d (proteomic data). Thus, Fig. 6c shows a pronounced red peak for cluster 3, reflecting its mixed-age composition, whereas a similar pattern is absent in Fig. 6d. The PCA results, performed separately for clinical laboratory (Fig. 6e) and proteomic (Fig. 6f) data, also demonstrate a lack of clear separation of samples by age categories, underscoring the complexity and continuum of the processes underlying aging.

To gain deeper insights into the heterogeneity of human aging, we performed a detailed analysis of cases that defied straightforward age-based clustering. Specifically, we examined two paradigmatic scenarios: (1) younger individuals grouped within clusters predominantly composed of older adults (segment 1.1., Fig. 6b), and (2) long-lived individuals clustered primarily with younger adults (segment 4.1., Fig. 6b).

To elucidate the molecular basis of these groupings, we compared the abundance of 15 key proteins between these age-incongruent groups. These proteins were selected based on their maximal contribution and predictive power within our previously defined proteomic clocks of aging (Fig. 3b), representing core biomarkers of the aging process itself.

Radial (spider) plots visualizing the multi-protein signatures of representative samples from each segment are presented in Fig. 7a and 7b.

Despite the stark contrast in the chronological age of the individuals compared, their profiles for these aging-related proteins show remarkable convergence, with no significant differences detected. This visual finding is substantiated by the detailed distribution of each individual protein within the analyzed groups (segment 1.1 in Fig. 7c and segment 4.1 in Fig. 7d).

Based on the analysis of a specific predefined panel of proteomic markers, these data indicatethe existence of distinct aging trajectories at the individual level. The observed patterns demonstrate that in some individuals, molecular programs associated with aging can be prematurely activated, resulting in a proteomic profile akin to accelerated aging. Conversely, others maintain a more youthful molecular phenotype despite advanced chronological age, indicative of decelerated or successful aging. Dissociation between chronological and molecular age underscores the critical limitation of using time-based metrics alone to assess biological state. Our findings highlight the presence of both resilience and vulnerability factors within age groups, suggesting that future research and potential interventions should target specific biological pathways rather than broad chronological categories. This personalized, phenotype-driven approach is essential for advancing a more precise understanding of aging and developing tailored strategies to promote healthspan.

Discussion

The association between routine clinical blood parameters, the proteomic profile, and the aging process was investigated in healthy individuals. Notably, a number of clinical laboratory parameters (somatomedin C (IGF-1), DHEA-S, cystatin C, urea, HbA1c) and proteins (cystatin C, complement C5, fibronectin, sex hormone-binding globulin, platelet basic protein (CXCL7), haptoglobin, ceruloplasmin, apolipoprotein B-100, beta-2-microglobulin) exhibit significant age-related changes and can be considered as potential biomarkers of aging. However, for the majority of parameters, no significant correlation was found between group mean values and chronological age within WHO-established categories. The observed findings are consistent with our previous results, where some proteins detected in plasma samples of astronaut candidates within one age group demonstrated extremely high inter-individual variations[27]. In the present study we discovered overlapping proteomic and clinical profiles in some individuals from the extreme young and old age groups. These results highlight the substantial inter-individual heterogeneity of the aging process. One major limitation of this work should be clearly distinguished. Comparisons between individuals of different ages were made at a single time point and utilized a pre-defined proteomic (n = 49) and clinical laboratory panel (n = 182) on a one cohort of 99 individuals. It is therefore important to emphasize that the following discussion and conclusions are context-specific and should not be generalized beyond this particular dataset without further validation.

Assessing the heterogeneity of molecular parameters with age, we performed variability analysis using a balanced (by sample size) distribution across age groups throughout the adult lifespan. The general age-dependent decline in proteomic variability may indicate a convergence or reduced diversity of individual aging profiles at the molecular level. The peak in variability during mid-adulthood (ages 28–45) underscores the significance of individual trajectories. This life stage can be considered a "biological middle ground," where the molecular phenotype is less driven by the high metabolic intensity of youth or the progressive metabolic decline of old age. Contrary to expectation, the level of variation in this group was higher than in both younger and older groups, pointing to a spectrum of unique trajectories rather than a uniform intermediate state. Cluster analysis of both clinical parameters and proteins independently confirmed significant heterogeneity, grouping individuals into clusters that did not align with chronological age. Interestingly, one study identified a midlife crisis, characterized by a decline in well-being, which typically occurs around the late forties[28]. This period may subsequently influence an individual's aging trajectory.

Our results demonstrate a key limitation of the population-based approach. While group averages show some expected trends, these patterns disappear at the individual level. Population-based approaches not only poorly predict personal aging trajectories but also mask the true biological reality defined by profound inter-individual heterogeneity. Consequently, aging cannot be reduced to a function of chronological time. We acknowledge that the present study does not provide definitive answers regarding the universal nature of aging trajectories. Nevertheless, our findings support the notion that attempts to standardize aging at the population level by identifying a universal age-associated phenotype have fundamental limitations, dictated by the unique biological aging trajectory of each organism. Clearly, this concept requires longitudinal observations and further validation in independent cohorts, using broader biomarker panels and repeated measurements.

The shift from population-averaged aging trajectories to individual-specific aging patterns raises the fundamental question of how to define this personalized process. More broadly, our observations align with a wider conceptual framework, which was not directly interrogated in the present study, but which provides a valuable foundation for interpreting our findings and generating new hypotheses. Following this framework, we propose viewing aging through the lens of the depletion of adaptive potential and compensatory reserve, the internal resources of an organism that determine its resilience and are linked to the concept of frailty. The key challenge becomes the quantitative assessment of these reserves in a specific individual. This raises critical questions regarding which molecular parameters to measure, their temporal dynamics, the standardization of a physiological challenge, and the definition of an individual baseline. A model for calculating resilience could integrate multi-omics and immunological data, encompassing genetic, transcriptomic, proteomic, and metabolomic profiles.

Aging is not solely an intrinsic biological process but is significantly modulated by external sociotechnomedical systems[29]. These systems function as catalysts or inhibitors of aging trajectories. They create a civilizational buffer that supports a post-aging state, where the continuation of biological life becomes decoupled from the exhaustion of endogenous reserves. A lot of effort has been made on

the expansion of our knowledge of how the environment shapes the aging phenotype. Such exposures as environmental factors, lifestyle and diet are included in the concept of exposome. The exposome approach was proposed in 2005 and can be defined as “the cumulative measure of environmental influences and associated biological responses throughout the lifespan including exposures from the environment, diet, behavior, and endogenous processes”[30]. According to this concept, environmental quality of air, water, soil and food, urban or rural environment, climate and weather, social determinants (e.g., poor housing, neighbourhood, social stratification) play role in one’s susceptibility to exposures and should be considered in aging research. Moreover, examples of the intersection between environmental stressors (air pollution, heavy metals, pesticides and others), aging biomarkers and aging related phenotypes has also been demonstrated[31]. These results highlight the complex interplay between external environmental and internal biological factors that influence an individual’s aging trajectory. In light of this, it seems insufficient to study markers in isolation to the environment, raising the question of which metadata to consider at the research design stage. Each person is characterized by their unique lifestyle, and the surrounding environment also differs across different life stages. For instance, individuals living in urban environments generally tend to experience chronic stress more often than those living in rural areas[32]. On a physiological level, prolonged activation of the hypothalamic-pituitary-adrenal axis—a key component of the stress system—leads to glucocorticoid-induced suppression of growth hormone secretion and inhibition of somatomedin C (IGF-1) effects on target tissues [33]. Consequently, this may lead to variations in IGF-1 levels among individuals of the same age and sex group who experience different environments. This example highlights the limitations of seeking "universal aging markers" without accounting for lifestyle factors. While the broader framework outlined above is speculative at this stage, it provides a useful interpretative context for our findings and highlights the need for more comprehensive data collection in future studies. We suggest that expanding metadata to include individual life and behavioral characteristics could lead to a more complete picture of unique aging characteristics.

Additional analysis confirms the complexity and multifaceted nature of the process: the identified sex differences in the stable core signature highlight distinct molecular axes. In men, complement C5 (C5) and SERPING1 consistently ranked among the top predictors, pointing to complement activation and protease regulation as dominant features of male aging. In women, retinol-binding protein 4 (RBP4), transthyretin (TTR), and sex hormone-binding globulin (SHBG) formed a consistent cluster, indicating shifts in metabolic transport and endocrine availability. These sex-specific patterns underscore the molecular heterogeneity of aging and reinforce the need for personalised approaches. Accordingly, a central principle that emerges from this work is personalization: the assessment of any changes relative to an individual baseline, where the "norm" of adaptive potential is a unique metric for each person.

Conclusion

The integrated analysis of the plasma proteome and clinical blood parameters in a healthy cohort (n = 99) reveals a fundamental characteristic of human biological aging: its profound inter-individual

heterogeneity. While we identify population-level trends based on the predefined proteomic (n = 49) and clinical laboratory (n = 182) panels and define a reproducible eight-marker signature indicative of systemic shifts in inflammation, complement activation, and tissue remodeling, these averages mask the true biological reality at the individual level. Our core finding, within these specific panels and a cohort, is the pronounced inter-individual heterogeneity in molecular profiles, which argues against a single universal trajectory tightly coupled to chronological age. The discovery of overlapping proteomic and clinical profiles in individuals from extreme age groups, coupled with age-dependent peaks in variability (in proteomics during mid-adulthood and in clinical parameters later in life), demonstrates that aging is not a uniform, linear process.

Our results underscore a critical limitation of population-centric "aging clocks." While useful for identifying group-level associations, such models are inherently poor predictors of individual aging paths and risk. They fail to capture the essential dynamics of an individual's adaptive potential and compensatory reserve as the internal resources that determine resilience and frailty. Consequently, aging cannot be validly reduced to a simple function of time.

Therefore, we propose a paradigm shift from a population-based to a person-centered framework. The future of aging research and intervention lies in quantifying an individual's physiological resilience. This requires moving beyond static biomarker measurements to dynamic assessments of the response to standardized challenges, integrating multi-omics data to model personal "resilience calculators." Furthermore, our data highlight the significant modulating roles of sex-specific biology and external, civilization-derived factors that create a "post-aging" buffer, extending lifespan beyond intrinsic biological limits.

In conclusion, within the limits of the biomarker panels and cohort examined, there is no single universal molecular signature of aging, but rather a complex interplay of molecular and physiological changes interpreted through the unique context of each individual's life. Embracing this heterogeneity is the key to developing meaningful strategies for personalized healthspan extension.

Methods

Demographic characteristics of the cohort

The study included 99 participants (49 females and 50 males, aged 18–90 years). They were enrolled according to strict inclusion criteria: age ≥ 18 years, absence of acute or chronic diseases (including infectious, oncological, and severe non-infectious conditions), and provision of signed informed consent. A complete list of inclusion and exclusion criteria, with justification, is provided in Table 1.

Samples were collected as part of the multicenter research project “RUSS-AGE” (protocol VA06/2022), which received approval from the Local Ethics Committee of the Russian Gerontological Scientific and Clinical Center (protocol No. 59 dated September 13, 2022). The Center operates as a separate structural unit of the Pirogov Russian National Research Medical University (RNRMU) under the Russian Ministry of

Health. All biomaterials and corresponding data were collected at the clinical facilities of this Center by its staff members.

Participant examination, sample collection, and biobanking

Eligible participants underwent questionnaire-based assessment, anthropometric measurements, evaluation of physical and cognitive status, and venous blood collection for subsequent proteomic analysis.

Blood sampling was performed by trained personnel using vacuum collection systems (K_2-EDTA tubes) strictly in the fasting state to minimise pre-analytical variability. Plasma was obtained by centrifugation (3000 rpm, 10 min), aliquoted, and stored at − 80°C in accordance with ISBER (International Society for Biological and Environmental Repositories) guidelines until transfer to the coordinating centre. For a comprehensive assessment of age-associated changes, 182 clinical and laboratory parameters were analysed, including standard biochemical markers, complete blood count indices, and hormonal markers. Detailed demographic and clinical characteristics of the cohort are provided in Table S1 - Supplementary.

Table 1
Inclusion and exclusion criteria for study participants.

Inclusion criteria	**1. Written informed consent to participate in the study.** **2. Age ≥ 18 years.**
Exclusion criteria	1. Refusal to participate in the study or to provide informed consent. 2. Presence of infectious diseases (hepatitis C, hepatitis B, including HBsAg carriage, HIV infection). 3. Presence of acute disease, exacerbation of chronic inflammatory disease, or surgical intervention within one month prior to enrolment. 4. Lack of remission of oncological disease or receipt of anticancer therapy less than three months before enrolment. 5. Severe cognitive or sensory impairment and psychiatric disorders that hinder contact with the participant. 6. Severe chronic non-infectious diseases (life-threatening arrhythmias, unstable angina, chronic heart failure NYHA class III–IV, chronic kidney disease stages 3b–5, type 1 diabetes, type 2 diabetes with terminal complications, systemic connective tissue diseases, osteoarthritis stage IV, as well as prior stroke, myocardial infarction, or low-energy fractures). 7. Pregnancy or lactation. 8. Any other factors deemed by the investigator to interfere with participation in the study.

Definition of Age Groups

Participants were stratified by age using two distinct schemes: the standard WHO classification and a balanced grouping to ensure more uniform sample sizes across the adult lifespan (see Table 2 for cohort details).

Table 2
Description of Study Groups According to Age Classifications.

Age Stratification	Age Range (years)	n	Mean Age ± SD (years)
WHO Classification	18–44	49	25.6 ± 6.1
	45–59	15	49.9 ± 4.4
	60–74	14	65.6 ± 5.4
	75–90	21	82.0 ± 3.0
Balanced Groups Classification	18–22	19	20.4 ± 1.3
	23–27	16	24.9 ± 1.7
	28–45	17	35.6 ± 6.0
	46–60	16	53.4 ± 5.2
	61–80	16	72.1 ± 6.8
	81–90	15	83.1 ± 2.8
Total Cohort	18–90	99	46.2 ± 18.7

Comparative analyses were performed using two distinct age stratifications: the four-category World Health Organization (WHO) classification and a novel balanced six-group scheme (detailed in Table 2). The balanced six-group structure was specifically selected as it optimized sample distribution, yielding the most uniform participant counts across cohorts (Fig. 8). This grouping strategy also minimized the pooled standard deviation of mean ages across groups (SD = 3.91) compared to alternative five-group (SD = 4.02) or seven-group (SD = 4.18) partitions, thereby reducing within-group age heterogeneity and enhancing analytical precision.

This grouping strategy also minimized the pooled standard deviation of mean ages across groups (SD = 3.91) compared to alternative five-group (SD = 4.02) or seven-group (SD = 4.18) partitions, thereby reducing within-group age heterogeneity and enhancing analytical precision.

Targeted proteomic analysis (SRM)

Proteomic profiling of samples by Selected Reaction Monitoring with stable isotope-labelled peptide standards (SRM/SIS) was conducted in accordance with a previously published protocol[34].

Briefly, proteins approved by the FDA for diagnostic use[34] were selected for quantification, as they serve as indicators of systemic dysfunction in molecular pathways that may contribute to biological ageing.

The SRM/SIS panel included 151 proteotypic peptides mapping to 111 proteins, selected based on the NextProt database (https://www.nextprot.org).

Each sample was measured in two technical replicates. Chromatographic separation was performed on an Agilent 1200 series system (Agilent Technologies) coupled to a TSQ Quantiva triple quadrupole mass spectrometer (Thermo Fisher Scientific). Peptides were separated using a ZORBAX SB-C18 column (150 × 0.5 mm, 5 µm) with a linear acetonitrile gradient at a flow rate of 20 µL/min, with prior column conditioning using 5% (v/v) acetonitrile in 0.1% (v/v) formic acid. Mass spectrometric detection was performed in dynamic MRM (dMRM) mode with Q1/Q3 isolation windows of 0.7 Da and a scan cycle time of 1.2 s, monitoring 2–4 transitions for each target protein.

Quantitative analysis was performed in Skyline (v4.1.0) by comparing peak area ratios of native versus SIS peptides (Table S3). Protein concentrations (nmol) were calculated as the mean of two replicates. Data processing and quantitative calculation of concentrations were carried out in Skyline (v4.1.0), taking into account all monitored transitions.

In total, SRM analysis yielded reliable signals for 68 peptides mapping to 49 proteins (Table S1 - Supplementary).

Data are available via ProteomeXchange with identifier PXD075101.

Statistical analysis

Biological age was calculated using the PhenoAge algorithm[35], based on a panel of clinical biomarkers: Albumin (g/dL), Alkaline Phosphatase (U/L), Creatinine (mg/dL), Glucose (mg/dL), C-reactive protein (mg/L), Lymphocytes (%), Mean Cell Volume (fL), Red Cell Distribution Width (%), and White Blood Cells (10^3/µL).

Further statistical analyses of clinical, biochemical, and proteomic data were performed in Python (v3.13.2) using the stats (v0.19.2), pandas (v1.5.3), and numpy (v1.26.0) packages. Data visualisation was conducted using Seaborn (v0.12.2) and Matplotlib.pyplot (v3.8.0). Pearson's test was used to assess correlations, the Shapiro–Wilk test for normality, and Levene's test for homogeneity of variance. Depending on the distribution of variables, comparisons were performed using ANOVA or the non-parametric Kruskal–Wallis test. For multiple testing correction, p-values were adjusted by the Benjamini–Hochberg procedure, with FDR < 0.05 considered statistically significant. Post hoc analysis was carried out with Dunn's test for multiple comparisons.

Regression analysis and feature selection

To identify key predictors of chronological age, we developed a linear regression model. Biomarkers with the highest absolute coefficient values in the regression model were selected as having the greatest influence on age prediction. Additionally, features were ranked based on their value dispersion (variability) to identify parameters with the highest individual heterogeneity, which may be informative for detecting distinct aging trajectories and subtypes (Fig. S1).

Statistical Definitions of Variability Metrics

To precisely characterize the observed heterogeneity, we employed two complementary measures of variability: the coefficient of variation (CV) and the variance.

The coefficient of variation ($CV = \sigma/\mu$), defined as the ratio of the standard deviation (σ) to the mean (μ), provides a normalized, scale-independent measure of dispersion. It represents an averaged, population-level statistic that quantifies relative variability, allowing for comparison across different biomarkers with distinct units. Its calculation inherently involves normalization by the sample mean.

In contrast, variance (σ^2) measures the average of the squared deviations from the mean, preserving the absolute scale of the data. Crucially, while it is also normalized by the sample size (n) in its standard computation [$\sigma^2 = \Sigma(x_i - \mu)^2/(n-1)$], it retains the squared contribution of each individual data point. This makes it sensitive to extreme values and reflects the total heterogeneity within a cohort, including the distinct contribution of individual outliers or specific subpopulations.

Hierarchical clustering was performed using Ward's linkage method with Euclidean distance metric on protein expression profiles and clinical parameters separately. Missing values were imputed using median substitution, and clustering was conducted with 3 clusters for protein markers and 3 clusters for clinical parameters based on both statistical and biological considerations.

Visualization and validation included dendrogram construction with age-labeled leaves color-coded by age intervals, supplemented by Principal Component Analysis (PCA) to visualize cluster separation.

Machine learning models

Prior to modelling, plasma protein concentrations were log-transformed to normalize their distributions, and ratios of the ten most correlated protein pairs were generated to capture potential biological interactions. All missing values in the combined feature set (49 proteins, 10 ratios, and 182 routine clinical laboratory parameters) were imputed with the median. Data were then scaled using StandardScaler.

For predictive modelling, we employed a Gaussian Process Regressor (GPR) with a radial basis function (RBF) kernel (ConstantKernel(1.0) * RBF(length_scale = 1.0)), implemented in scikit-learn (v1.3.0). The kernel hyperparameters were optimized by maximizing the log-marginal likelihood, and the regularization parameter α was set to 0.1. To evaluate generalizability, we used a nested cross-validation scheme with 5 outer folds and 5 inner folds. Within each training fold, we performed median imputation (if needed), variance threshold filtering (threshold = 0.01), scaling with StandardScaler, and feature selection using SelectKBest with F-regression. The number of selected features (k) was set to 10 for the combined cohort and men, and to 5 for women, based on empirical performance. All preprocessing steps were performed strictly inside each training fold to prevent data leakage. Model performance was assessed on the held-out test fold using Pearson correlation (r), coefficient of determination (R^2), and mean

absolute error (MAE), averaged across the five outer folds. Confidence intervals (95%) were estimated by bootstrapping the fold-wise metrics.

To identify robust and reproducible biomarkers, we additionally evaluated the stability of feature selection using a bootstrap resampling procedure. For each cohort (combined, men, women), we generated 60 bootstrap samples by resampling donors with replacement, preserving the original cohort size. On each bootstrap sample, we trained both Random Forest[36] and XGBoost[37] models and assessed feature importance via permutation testing on the out-of-bag samples, which avoids in-sample bias. The top-8 features (ranked by importance) were recorded for each bootstrap iteration, and the selection frequency of each feature was computed as the proportion of iterations in which it appeared in the top-8. Frequencies from the two algorithms were averaged, and features with a mean selection frequency ≥ 75% were considered candidates for a stable core signature. For comparison, Random Forest and XGBoost were also trained with default hyperparameters using the same nested cross-validation framework[38]; feature importance was evaluated via permutation testing[39]. Full results for all models are provided in Table S4.

Principal component analysis (PCA) with two components was applied to the preprocessed data for visualization purposes, with results coloured by age group and sex.

All analyses were implemented in Python (v3.13) using the scikit-learn, pandas, numpy, and matplotlib libraries. Code and data are available upon request.

Declarations

Funding Declaration

This research was financially supported by the Ministry of Science and Higher Education of the Russian Federation within the framework of Agreement No. 075-15-2024-643.

Author Contributions

Conceptualization, E.V.S., V.V.B., A.A.K., A.I.G., A.S.K., S.N.T., E.V.I., E.A.P. and V.M.G.; methodology, E.V.S., E.A.P. and A.A.K.; software, E.V.S. and A.Ab.; validation, V.V.B., A.A.K., A.I.G. and A.S.K.; formal analysis, E.V.S., A.Ab., A.A.G. and A.I.G.; investigation, S.N.N. and V.G.Z.; resources, I.D.S., A.V.C. and O.N.T.; data curation, E.V.S. and A.I.G.; writing—original draft preparation, E.V.S., V.V.B., E.A.P. and A.I.G.; writing—review and editing, E.V.S., V.V.B., E.A.P., A.V.L., V.M.G., S.N.T. and A.I.G.; visualization, E.V.S. and A.I.G.; supervision, E.A.P., E.V.I. and A.I.A.; project administration, E.A.P.; funding acquisition, E.A.P. All authors have read and agreed to the published version of the manuscript.

Competing Interests

The authors declare no competing interests.

Data and Code Availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[40] partner repository with the dataset identifier PXD075101.

References

1. Guo, J. *et al.* Aging and aging-related diseases: from molecular mechanisms to interventions and treatments. *Signal Transduction and Targeted Therapy 2022 7:1* 7, 391- (2022).
2. Tenchov, R., Sasso, J. M., Wang, X. & Zhou, Q. A. Aging Hallmarks and Progression and Age-Related Diseases: A Landscape View of Research Advancement. *ACS Chem Neurosci* 15, 1–30 (2023).
3. Teschendorff, A. E. & Horvath, S. Epigenetic ageing clocks: statistical methods and emerging computational challenges. *Nature Reviews Genetics 2025 26:5* 26, 350–368 (2025).
4. Min, M., Egli, C., Dulai, A. S. & Sivamani, R. K. Critical review of aging clocks and factors that may influence the pace of aging. *Frontiers in Aging* 5, 1487260 (2024).
5. Huang, H. *et al.* Decoding aging clocks: New insights from metabolomics. *Cell Metab* 37, 34–58 (2025).
6. Tao, X. *et al.* Biomarkers of Aging and Relevant Evaluation Techniques: A Comprehensive Review. *Aging Dis* 15, 977–1005 (2024).
7. Zha, K. *et al.* Circadian Rhythm: Biological Functions, Diseases, and Therapeutic Targets. *MedComm (Beijing)* 6, e70435 (2025).
8. Miller, M. *et al.* Impact of acute stress exposure on genome-wide DNA methylation. *Scientific Reports 2025 15:1* 15, 23931- (2025).
9. Lu, W. H. Effect of Modifiable Lifestyle Factors on Biological Aging. *JAR life* 13, 88 (2024).
10. Fierro-Marrero, J., Reina-Varona, Á., Paris-Alemany, A. & La Touche, R. Frailty in Geriatrics: A Critical Review with Content Analysis of Instruments, Overlapping Constructs, and Challenges in Diagnosis and Prognostic Precision. *Journal of Clinical Medicine 2025, Vol. 14, Page 1808* 14, 1808 (2025).
11. Deng, Y. T. *et al.* Atlas of the plasma proteome in health and disease in 53,026 adults. *Cell* 188, 253–271.e7 (2025).
12. Blume, J. E. *et al.* Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. *Nat Commun* 11, (2020).
13. Shen, X. *et al.* Nonlinear dynamics of multi-omics profiles during human aging. *Nature Aging 2024 4:11* 4, 1619–1634 (2024).
14. Jin, Z. *et al.* Decoding sexually dimorphic proteomic landscapes in the context of aging and mortality. *Communications Medicine 2025 5:1* 5, 403- (2025).
15. Lehallier, B. *et al.* Undulating changes in human plasma proteome profiles across the lifespan. *Nat Med* 25, 1843–1850 (2019).
16. Ló Pez-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. Leading Edge Hallmarks of aging: An expanding universe. *Cell* 186, 243–278 (2023).

17. Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging* 10, 573–591 (2018).
18. Toth, L. *et al.* Age-related decline in circulating IGF-1 associates with impaired neurovascular coupling responses in older adults. *Geroscience* 44, 2771–2783 (2022).
19. Odden, M. C. *et al.* Age and cystatin C in healthy adults: a collaborative study. *Nephrol Dial Transplant* 25, 463–469 (2010).
20. Chen, X. xin *et al.* Blood urea nitrogen to creatinine ratio is associated with physical frailty in older-aged Chinese: a cross-sectional study. *Aging Clin Exp Res* 35, 581–589 (2023).
21. Samaras, N., Samaras, D., Frangos, E., Forster, A. & Philippe, J. A review of age-related dehydroepiandrosterone decline and its association with well-known geriatric syndromes: is treatment beneficial? *Rejuvenation Res* 16, 285–294 (2013).
22. Deutsch, E. W. *et al.* Human Proteome Project Mass Spectrometry Data Interpretation Guidelines 3.0. *J Proteome Res* 18, 4108–4116 (2019).
23. Tammen, H. *et al.* Prerequisites for peptidomic analysis of blood samples: II. Analysis of human plasma after oral glucose challenge -- a proof of concept. *Comb Chem High Throughput Screen* 8, 735–741 (2005).
24. Wang, S. J. *et al.* Attribution of the discrepancy between ELISA and LC-MS/MS assay results of a PEGylated scaffold protein in post-dose monkey plasma samples due to the presence of anti-drug antibodies. *Anal Bioanal Chem* 402, 1229–1239 (2012).
25. Yang, M., Harrison, B. R. & Promislow, D. E. L. Cellular age explains variation in age-related cell-to-cell transcriptome variability. *Genome Res* 33, 1906–1916 (2023).
26. Viñuela, A. *et al.* Age-dependent changes in mean and variance of gene expression across tissues in a twin cohort. *Hum Mol Genet* 27, 732–741 (2018).
27. Kopylov, A. T. *et al.* Targeted Quantitative Screening of Chromosome 18 Encoded Proteome in Plasma Samples of Astronaut Candidates. *J Proteome Res* 15, 4039–4046 (2016).
28. Blanchflower, D. G. Unhappiness and age. *J Econ Behav Organ* 176, 461–488 (2020).
29. Navarro, C. *et al.* Intrinsic and environmental basis of aging: A narrative review. *Heliyon* 9, e18239 (2023).
30. Wild, C. P. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol. Biomarkers Prev.* 14, 1847–1850 (2005).
31. Malecki, K. M. C. et al. Integrating environment and aging research: opportunities for synergy and acceleration. *Front. Aging Neurosci.* 14, 824921 (2022).
32. Xu, J. et al. Effects of urban living environments on mental health in adults. *Nat. Med.* 29, 1456–1467 (2023).
33. Chrousos, G. P. Organization and integration of the endocrine system. *Sleep Med. Clin.* 2, 125–145 (2007).

34. Novikova, S. E. *et al.* Mass-Spectrometric MRM Analysis of FDA-Approved Proteins in Plasma of Healthy Volunteers. *Biochem Mosc Suppl B Biomed Chem* 15, 40–61 (2021).
35. Liu, Z. *et al.* A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: A cohort study. *PLoS Med* 15, e1002718–e1002718 (2018).
36. Breiman, L. Random forests. *Mach Learn* 45, 5–32 (2001).
37. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 13-17-August-2016, 785–794 (2016).
38. Altmann, A., Toloşi, L., Sander, O. & Lengauer, T. Permutation importance: a corrected feature importance measure. *Bioinformatics* 26, 1340–1347 (2010).
39. Arlot, S. & Celisse, A. A survey of cross-validation procedures for model selection. https://doi.org/10.1214/09-*SS054* 4, 40–79 (2010).
40. Perez-Riverol, Y. et al. The PRIDE database at 20 years: 2025 update. Nucleic Acids Res. 53, D543–D553 (2025).

Figures

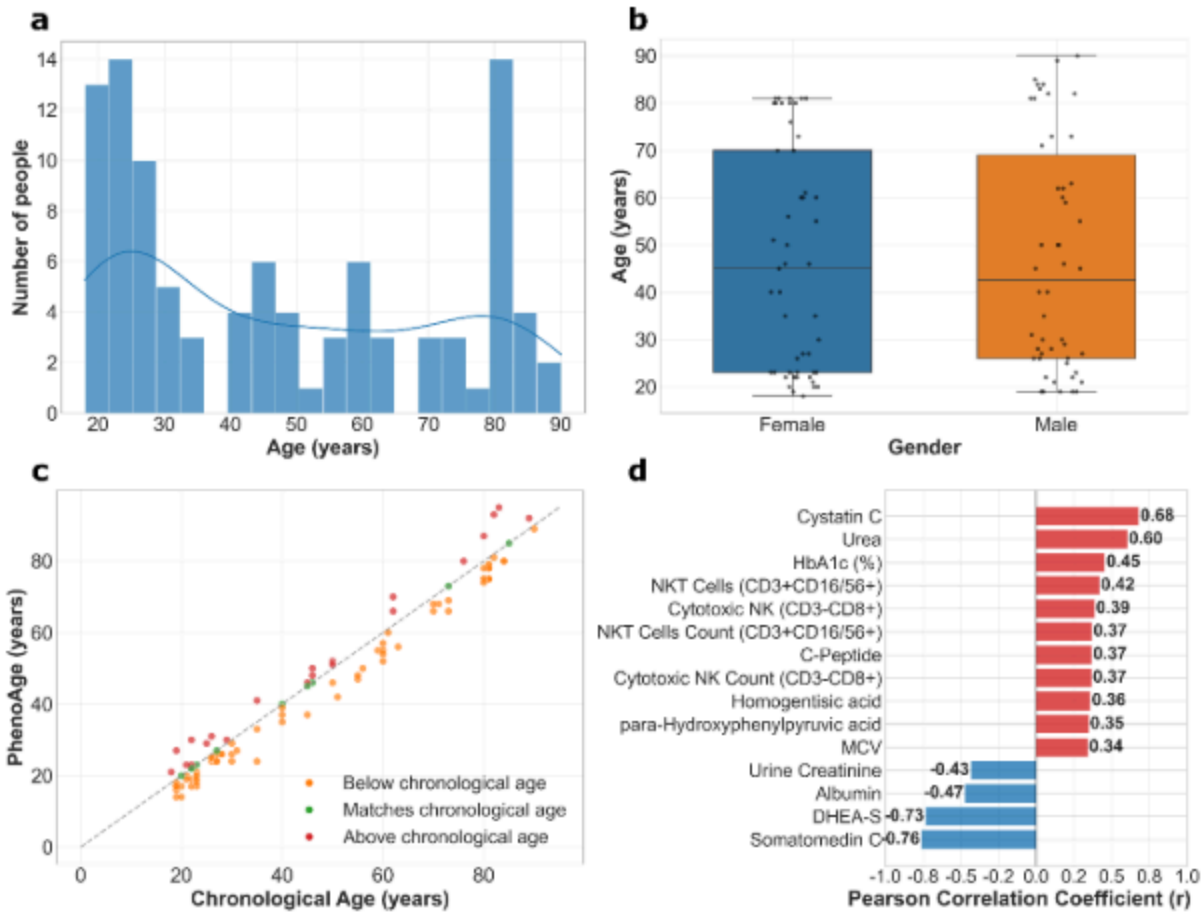


Figure 1

Characteristics of the study cohort. a, Distribution of participants by chronological age. **b,** Sex distribution (women, n = 49; men, n = 50). **c,** Relationship between chronological age and biological age predicted by the PhenoAge algorithm. **d,** Top 15 laboratory parameters significantly correlated with chronological age (Pearson's r, p < 0.05).

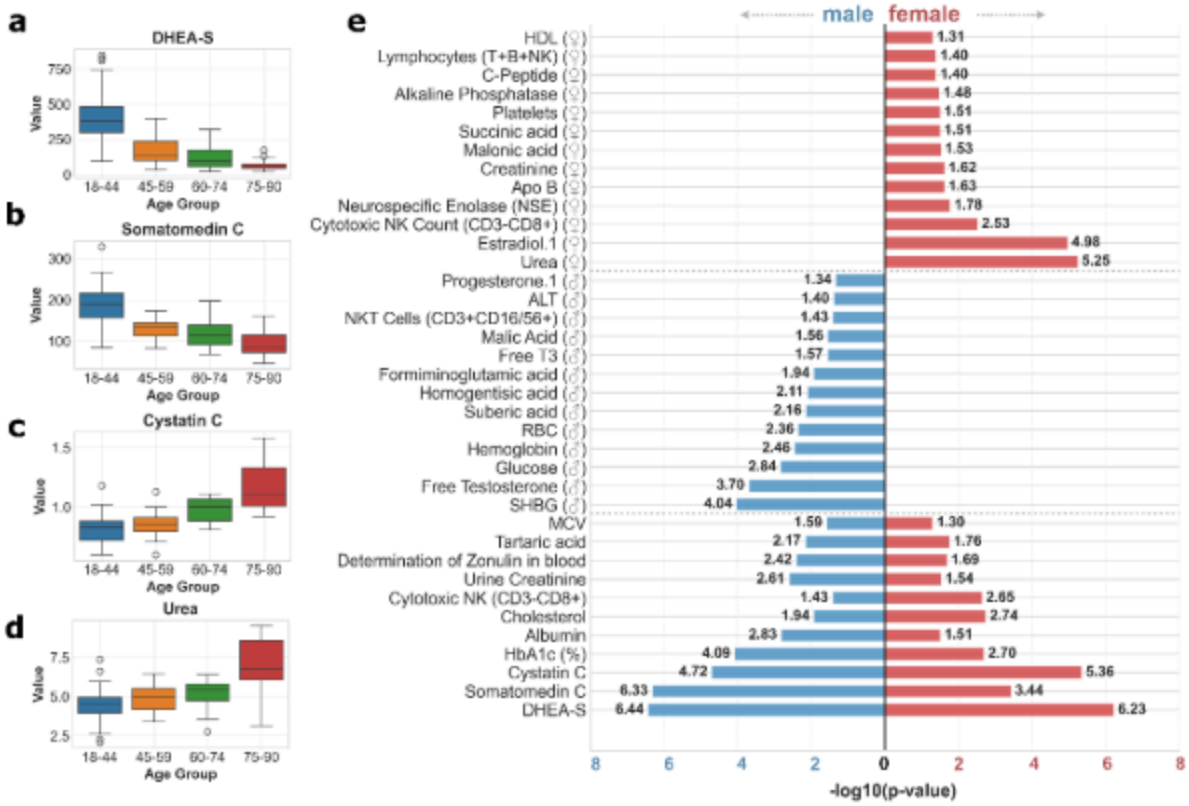


Figure 2

Age-associated alterations in clinical and biochemical parameters reveal sex-specific dynamics. **a**,**b**,**c**,**d,** Significant age-related differences in clinical and biochemical parameters (a-d) parameters with the highest statistical significance across age categories (ANOVA/Kruskal–Wallis, p < 0.05, followed by Tukey's post-hoc test. **e**, Distribution of significance scores (-log10(p-value)) for age-associated alterations in male and female datasets.

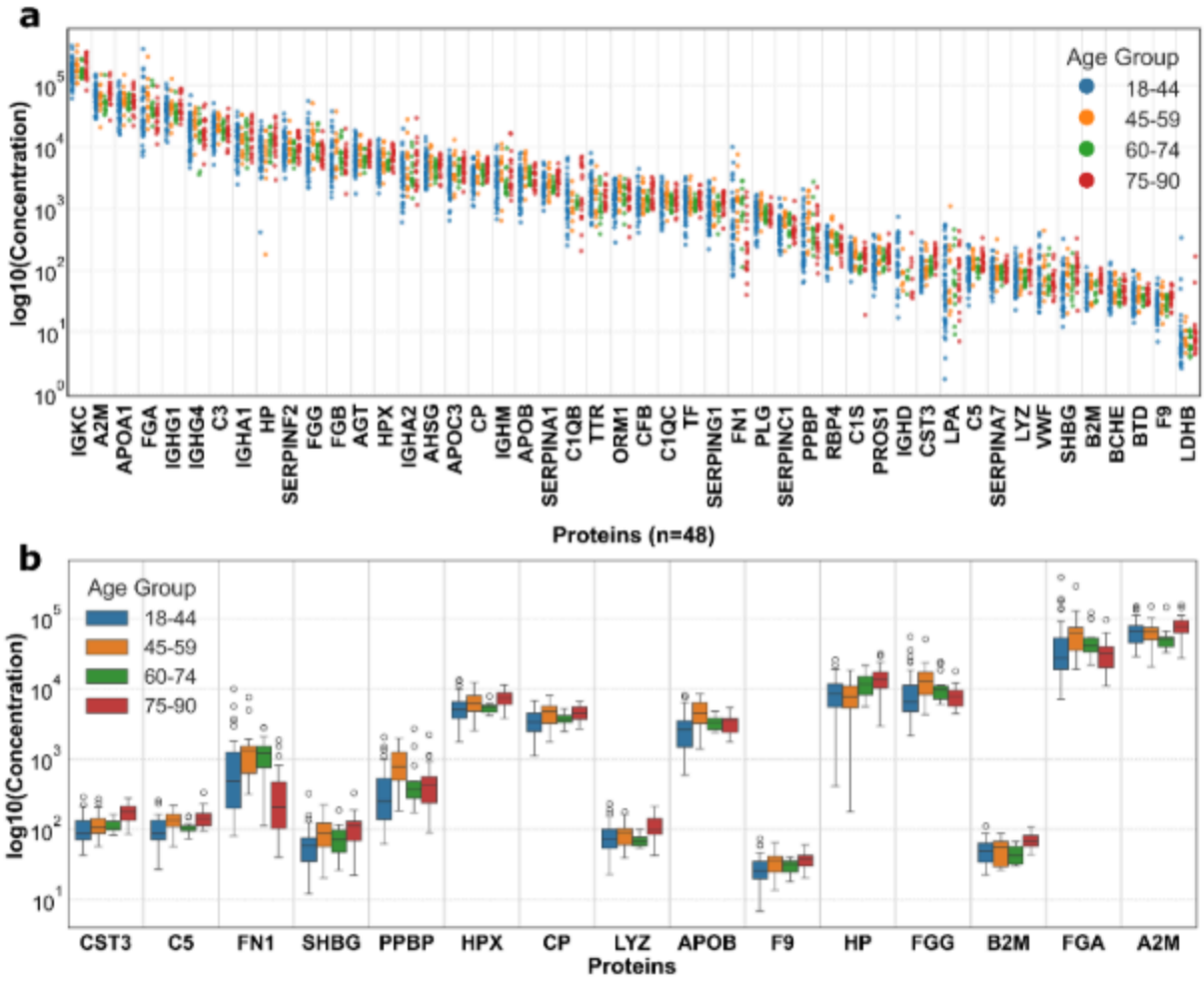


Figure 3

Targeted proteomics reveals age-associated dynamics of plasma protein concentrations. a, Distribution of protein concentrations (nM) measured by Selected Reaction Monitoring (SRM) across age groups classified according to WHO criteria. Each point represents an individual sample. **b**, Box plots of the proteins exhibiting the most significant age-related concentration changes (ANOVA or Kruskal–Wallis test, $p < 0.05$). The center line represents the median, box limits indicate the interquartile range (IQR), and whiskers extend to 1.5× IQR.

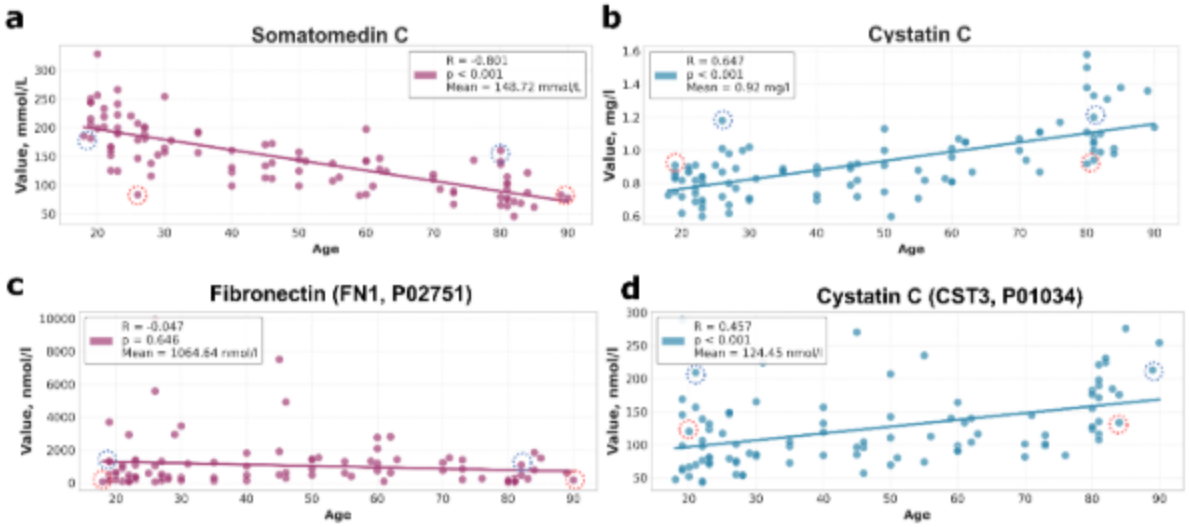


Figure 4

Divergent age-related trajectories of clinical and proteomic biomarkers reveal individual variability. a,b, Clinical biochemical parameters: (**a**) Serum Somatomedin C (IGF-1) concentration versus chronological age (n = 99), (**b**) Serum Cystatin C concentration versus age. Solid line, linear regression fit. **c**,**d**, Targeted proteomic (SRM) markers. (**c**) Plasma Fibronectin (FN1) concentration versus age. (**d**) Plasma Cystatin C (CST3) concentration versus age. Across all panels, dashed circle points highlight individual samples from opposite age extremes exhibiting similar biomarker levels.

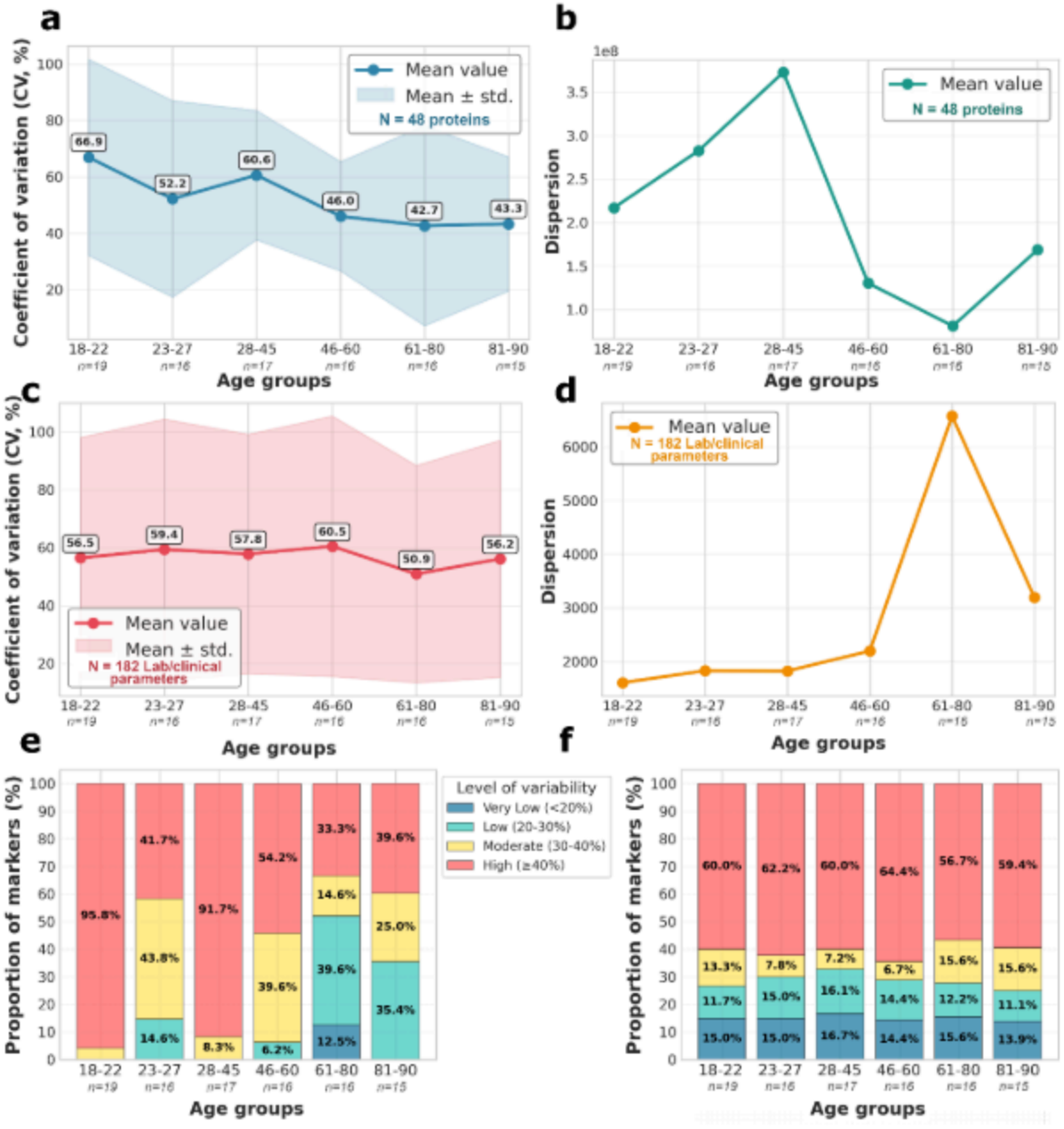


Figure 5

Age-dependent heterogeneity differs between the plasma proteome and clinical biochemical parameters. a,b, Age-group distribution of coefficient of variation (CV, %) and dispersion (IQR/median) for proteins measured by targeted proteomics (SRM). **c,d,** Corresponding variability metrics for routine clinical biochemical parameters. **e,f,** Stratification by variability level. Distribution of (**e**) proteomic and (**f**) clinical parameters across predefined CV categories (very low, low, moderate, high) within each age group.

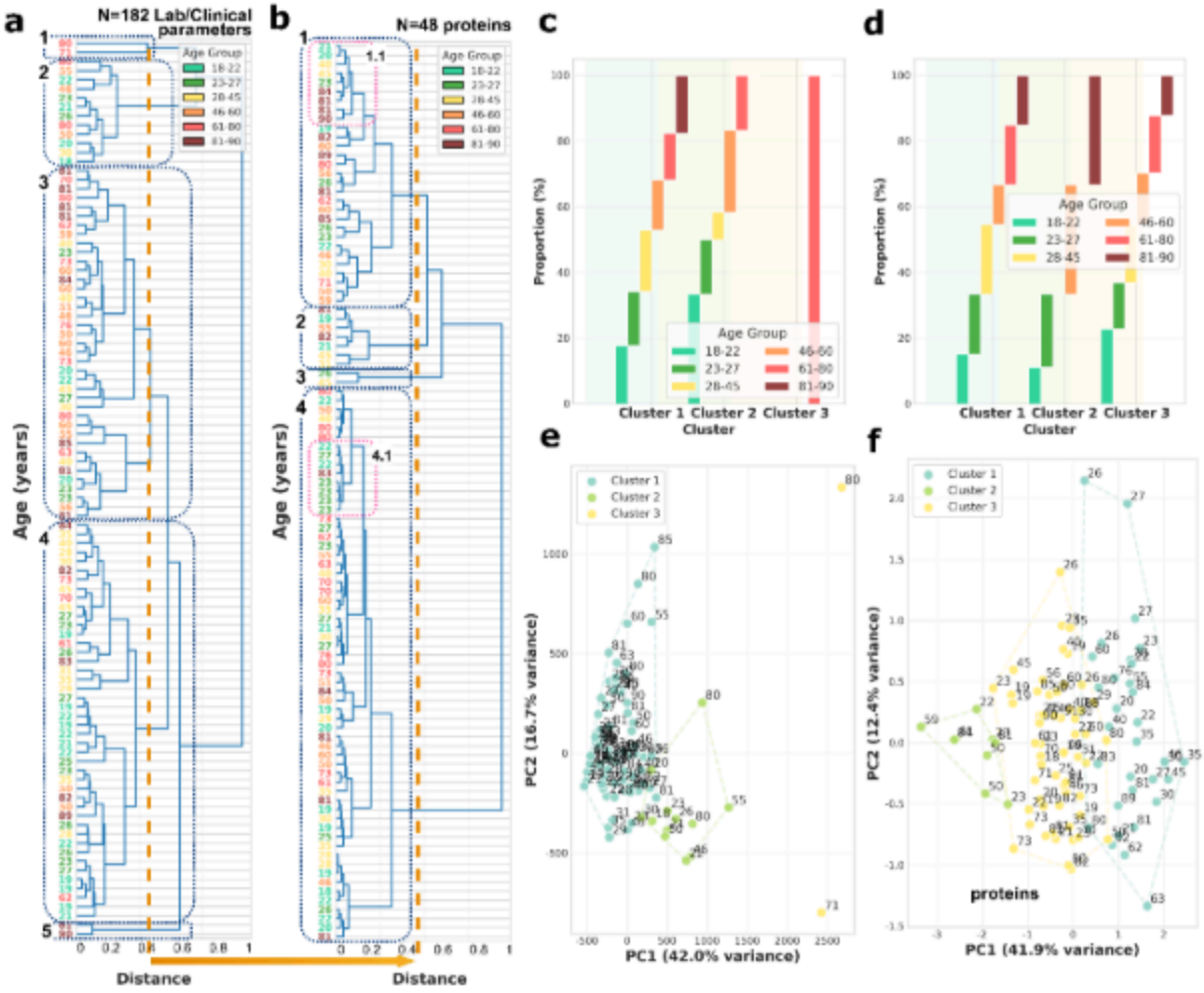


Figure 6

Integrated analysis identifies age-dependent biological signatures. **a**,**b**, Hierarchical clustering heatmaps. (a) Clinical laboratory parameters. (b) Targeted plasma proteome (SRM). Cluster numbers are indicated. **c**,**d**, Age distribution within clusters. Bar plots showing the proportion of individuals from each age group within the identified (**c**) clinical and (**d**) proteomic clusters. **e**,**f**, Principal component analysis (PCA) with age annotation. (**e**) PCA of clinical data, colored by cluster assignment. (**f**) PCA of proteomic data, colored by cluster assignment. Individual data points are labeled with the corresponding chronological age (years).

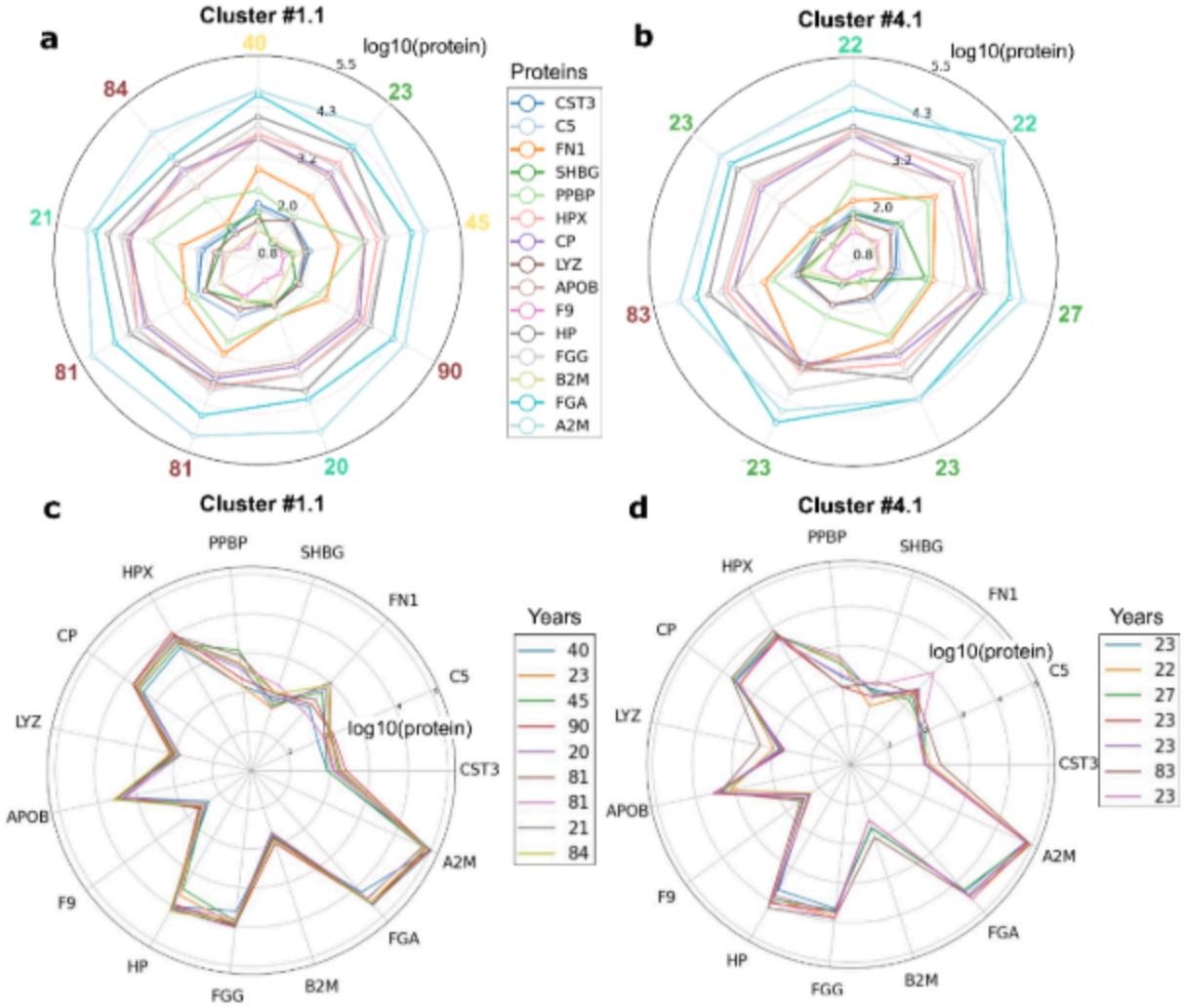


Figure 7

Radar visualization of aging-associated plasma proteome signatures. a,b, Individual-centric analysis. Radar plots depicting the plasma abundance profiles (log_{10}-transformed) of 15 aging-associated proteins. **c,d,** Protein-centric analysis. Cluster #1.1 ("younger phenotype within older cohort") (**a**,**c**), Cluster #4.1 (**b**,**d**) ("older phenotype within younger cohort"). All protein abundance values are log_{10}-transformed.

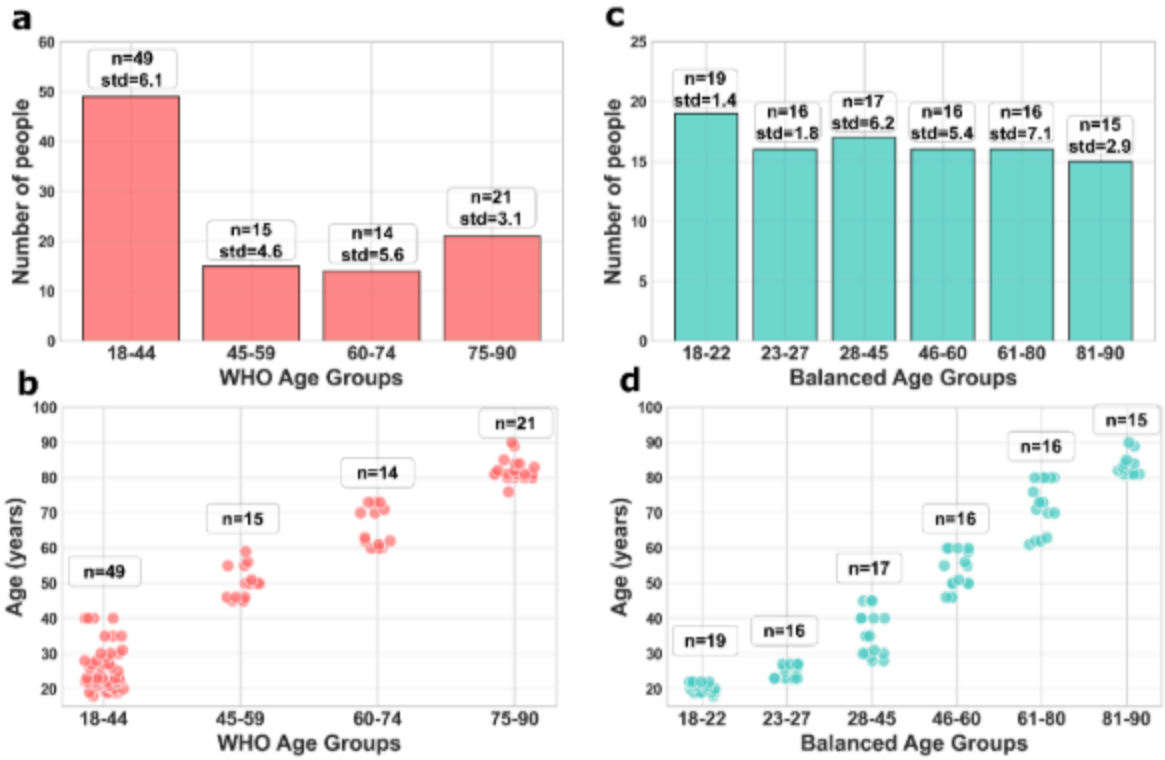


Figure 8

Participant distribution across the study's age groups.

This grouping strategy also minimized the pooled standard deviation of mean ages across groups (SD = 3.91) compared to alternative five-group (SD = 4.02) or seven-group (SD = 4.18) partitions, thereby reducing within-group age heterogeneity and enhancing analytical precision.

Supplementary Files

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

- SupplementaryTables.xlsx
- SupplementaryFigures.docx
- floatimage1.png